Journal of Experimental Psychology

VOLUME II, 1917

PUBLISHED BI-MONTHLY BY

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, N. J.

Entered as second-class matter at the post-office at Princeton, N. J.

Press of
The New Era Printing Company
Lancaster, Pa

CONTENTS OF VOLUME II

Journal of Experimental Psychology

Vol. II, No. 1. February, 1917

ON THE MEASUREMENT OF VISUAL STIMULATION INTENSITIES

BY LEONARD T. TROLAND

Nela Research Laboratory, National Lamp Works of General Electric Company, Nela Park, Cleveland, O.

I. Introduction

It must probably be admitted that there are few fields of science in which definite quantitative results are obtainable, which have been more carelessly cultivated than that of visual psycho-physiology. The literature of visual research is truly monumental, the ascertained qualitative facts are legion, and yet the laws of vision are few and vague. The conditioning cause of our present chaotic conception of visual response is undoubtedly the failure of the majority of investigators in this realm to pay attention to details, chiefly, their failure to measure in absolute units the conditions and results of their experiments. As a consequence, the conditions are not reproducible, and the results can be employed in the support of almost any hypothesis, at the experimenter's pleasure.

The essentials for the standardization and accurate description of the work in vision have been available to intelligent students of the subject for a century, but only quite recently have these essentials been developed to a form in which they are readily applicable. There are now a number of workers on vision in this country whose methods can be described as exact and scientific. Much of this development—on the photometric side at least—has been the out, come of the technical demands of illuminating engineering-

and the debate over many methodological details is still far from arriving at a definite conclusion.

The conditions of stimulation for any experiment in monocular vision are unequivocally determined if the 'energy distribution curve' (*cf.* Fig. 1) is known for the total radia-

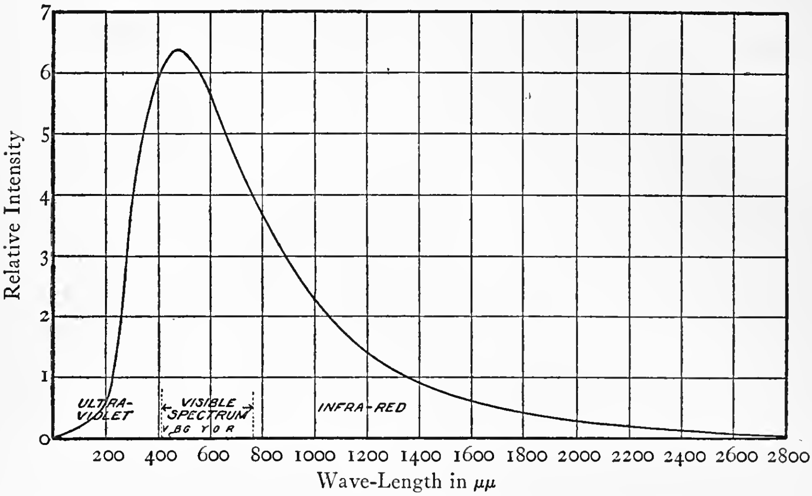


FIG. 1. Energy Distribution Curve of the Radiation from a "Black Body" at 6200° Absolute. The ordinates of this curve represent the 'specific radiant power density' for each wave-length, *i. e.*, they are proportional to the energy per unit wave-length, passing through a given surface in the path of the radiation, in a given time. The area enclosed between the curve and the axis of the abscissæ, is proportional to the 'total radiant power density.' (See note 1.) This distribution curve is closely similar to that of sunlight, and the sensation produced by such radiation would be approximately 'white.' It will be noticed that more than half of the energy lies outside of the visible spectrum. A thermopile will measure all of this energy, while a photometer can deal only with that lying in the visible spectrum.

tion falling upon each sensitive element of the retina. Since the visual receptors, proper, lie in the external strata of the retina, the state of affairs outside the eye is of importance only in so far as it determines that within the eye. The mode of this determination is not simple for any of the dimensions of the visual stimulus.

It is the purpose of the present article to give special consideration only to the general dimension of stimulation intensity. The discussion will be divided into three more or

less independent parts, the first dealing with the significance and usefulness of photometric as compared with radiometric measurements, the second considering the problem of the proper method for general photometry, and the third dealing with the influence of pupillary size upon visual stimulus intensity. The final object of the paper is the definition and justification of a standard unit of visual stimulus intensity, the *photon*. Little will be said which is essentially new, but in this field, at the present time, there is small danger that repetition will become over-emphasis.

II. Radiometry Versus Photometry in the Measurements of Stimulus Intensities

It no longer savors of originality to point out that the term 'intensity' has been employed in a very indefinite way by writers in psycho-physiological optics. It has been purposively so employed in the title of the present article. For a technical analysis of various possible meanings of the word, the reader is referred to a footnote,[1] which explains

[1] Rand (Psychol. Mongr., No. 62, 12) says: "This term [intensity] has been employed at various times to indicate (*a*) the energy of a beam of spectral light homogeneous as to color; (*b*) the white-value of a color; (*c*) the saturation of a color; and (*d*) the energy of light-waves reflected from a pigment surface as conditioned by the general illumination of the visual field." Further (p. 20) "The terminology which we propose to use . . . may be outlined as follows: *Intensity of stimulus* will be used to indicate the energy of light-waves coming to the eye. *Intensity of sensation*, or apparent intensity, will be used as its correlative subjective term. So used, it will signify merely energy or voluminousness of sensation and will have no reference whatever to the white-value of a color. . . . The terms *brightness* and *white-value* will be used interchangeably to indicate the lightness or darkness of a color."

Rand's system of intensity terms appears to be not less confused than that of previous workers. The fact is that practically every intensity term in vision has been employed at one time or another to denote the meanings of every other intensity term, so that any definite nomenclature must be more or less arbitrary. In the present paper the following classification will be utilized.

Intensity will be employed as a generic term to stand for any one of the group of allied dimensions which we are here discussing. The simplest method of defining these dimensions for the purposes of visual physiology is to consider them with direct reference to the retinal image.

From this point of view they are (1) the *total radiant power density*, or the number of ergs of radiant energy of all wave-lengths, striking the retina, per unit area, per unit time. This is the integral of the energy distribution curve for unit time and area, and is what would be measured by a surface thermopile (at the retina).

somewhat in detail the concepts to be used in the ensuing discussion.

One reason why the generalized visual stimulus is difficult to deal with lies in its extreme complexity. Physically speaking there is no such thing as 'homogeneous' light,[1] or a visual stimulus of a single wave-length. Homogeneity is a relative term only, since every light, no matter how 'pure,' must occupy a finite range of the spectrum, and hence must be constituted by an infinite number of wave-lengths, each

(2) *Specific radiant power density* or the number of ergs per unit wave-length of any single wave-length, striking the retina per unit area, per unit time. This is the value of the ordinate of the distribution curve for any given wave-length, or is the value of the derivative of the complete radiant power density with respect to the wave-length, for a given wave-length.

(3) *Retinal illumination*, or the luminous flux density at the retina, the number of *lumens* of *light* impinging upon the retina per unit area. This may be either *total* or *specific*, according as lights of all wave-lengths or of one wave-length, respectively, are considered. In the latter case the measure will be in lumens per unit wave-length.

(4) *Photometric brightness*, or the luminous intensity per unit projected area of any stimulus surface measured by the standard method of photometry, including only a surface of dimensions negligibly small in comparison with the distance to the observer. This, also, may be *total* or *specific* and, for a constant pupil, would be proportional to the retinal illumination, or *vice versa.* Photometric brightness is an external measure depending on the eye only for the relative values given to light of differing wave-lengths. It is measured in candles per unit area, or in lamberts.

(5) *Luminosity*, or apparent brightness; which is a wholly psychological variable, probably depending upon the degree of stimulation of the retina by given radiation. It cannot be expressed in any physical units, although equality of luminosities furnishes the basis for all photometric equations made by direct comparison. For a wide range of intensities, photometric brightness is independent of the absolute value of the retinal illumination, but the luminosity depends directly upon this quantity, and also upon the general state of sensitivity of the visual system.

(6) *Flicker value*, or the photometric brightness of any stimulus surface, as determined by the standard method of flicker photometry.

In order to avoid the introduction of odious technical terms into the text of the article, *total energy* will be employed as a synonym for 'total radiant power density,' and *specific* (*radiant*) *intensity* for 'specific radiant power density.' This amounts merely to a neglect of the fact that radiation has a definite energy density in space and is in motion.

A recent summary of photometric terminology is that of the 1915 Report of the Committee on Nomenclature and Standards of the Illuminating Engineering Society, *Trans. Illum. Eng. Soc.*, 1915, 10, 642–651. See also Rosa, E. B., 'Photometric Units and Nomenclature,' *Bull. Bur. of Stand.*, 1911, 6, 543–573.

[1] Exception might be taken to this statement on the basis of the modern 'quantum' theory of radiation, but such an exception is hardly relevant to the purposes of the present discussion.

with its own specific intensity. Only when the range is very short can we legitimately choose the average wave-length as representative of the whole range, and employ the total energy as the intensity measure. In all other cases we must either divide up the spectrum of the stimulus into a finite number of 'small ranges' of this sort, and state the total energy of each, or we must specify the function (or energy distribution curve) connecting specific (radiant) intensity with wave-length.

An important corollary of the above is that, physically, there is no such generic entity as 'white light,' a conception which is of so much importance in visual physiology. We might define white light physically as light, the energy distribution curve of which approximates that of solar radiation, or possibly that of the radiation from a so-called 'black body' at some definite temperature (see Fig. 1). However, the distribution curves of such types of radiation would be merely isolated examples, out of an infinite number of similar curves, having no essential peculiarities. A *uniform distribution* of energy over the visible spectrum would give rise to a sensation of (unsaturated) purple and not of white.

Another fertile cause of confusion in the discussion of the intensity relations of visual stimuli, is the double or compound meaning of the word 'light.'[1] Light, on the one hand, is a form of radiation, or moving electromagnetic energy, and on the other hand, is a quality of experience, or one dimension—at least—of visual sensation. According to current definitions, light-intensity—measured in lumens—is equal to radiant power—measured in watts—*multiplied by the relative luminosity producing capacity of the given radiation for the normal eye.*[2] This relative luminous capacity, 'stimulus coefficient,' or 'visibility,' is determined by a photometric equation of the given radiation to a standard, the radiant power of the standard being known or, at least, constant. Light, itself, is thus technically neither radiation nor visual

[1] *Cf.* Nutting, P. G., 'The Luminous Equivalent of Radiation,' *Bull. Bur. Stand.*, 5, 261–264, 1908. Also Cobb, P. W., 'Photometric Considerations Pertaining to Visual Stimuli,' *Psychol. Rev.*, 23, 72, 1916.

[2] *Cf.* Cobb, *loc. cit.*, pp. 87–88.

sensation but is a mathematical concept based upon both of these variables. When we speak of light we imply both radiation and sensation; but both radiation and visual luminosity can exist without any light existing. Light is measured by photometry, radiation by radiometry.

Any mass of radiation moving through space has a definite energy density—contains a definite number of ergs per cubic centimeter—and when this radiation falls upon a surface there is a definite flux of energy into that surface. The total intensity of such a flux on a unit area, can be completely specified in terms of ergs per second—or some other *unit of mechanical power* such as the watt. If the energy is wholly absorbed, and is converted only into heat, the power can be measured by calorimetric methods, *i. e.*, by ascertaining the rise in temperature of the absorbing body, the mass and specific heat of which are known. In practice, this is done, although not easily, with the help of a bolometer, a thermopile or a similar device. Determinations of this sort, when made for successive small wave-length ranges over the entire range of wave-lengths in the stimulus, would make possible an exact specification of the intensity, and would thus render the conditions of the experiment quite reproducible.

Two distinct, but closely related problems are involved in the control of visual stimulation intensities. The first is that of the equation of intensities, while the second concerns the establishment of a definite relation between at least one member of a set of such equations, and an *absolute* and reproducible *standard of intensity.*

In the study of visual response we are interested to determine the manner in which the various dimensions of the response depend upon each other and upon those of the stimulus. This dependency is complex in such a way that it can be represented symbolically by a polyvariable function, like $r = f(w, i, s, \cdots)$ where r is some one dimension of the response, and $w, i, s, \cdots$ are other dimensions, either of the response or of the stimulus. To determine the form of this complex function experimentally it is necessary to hold all of the variables on the right-hand side of the equation con-

stant, except one, and then to subject this one to variation. For example, to determine the effect of stimulus intensity, we may take any constant wave-length and try our experiment with different values of the intensity, or to find the influence of wave-length, we may select a definite intensity and vary the wave-length. In this latter procedure we are obliged to *equate* the intensities of the qualitatively different lights which we use.

For establishing such an equality of intensity the two general alternatives of photometry and radiometry are open. If the first method is employed the equation will be one of brightness; with the second method it will be one of radiant power (per unit area). The specific interpretation which is made of the term intensity will thus depend upon the method adopted.

The relation between the two types of equations should be borne carefully in mind. In the first place, when all other conditions are the same for the two equated stimuli—*i. e.*, for the same wave-length constitution, state of adaptation, position in the visual field, contrast, etc.—a photometric equation implies a radiometric equation. On the other hand, when the two stimuli are not similarly conditioned, in general this implication will not hold, and when the difference is one of wave-length a photometric equation may entail a relation between the radiometric intensity of one stimulus and that of the other, such that the quotient of these two intensities can have values ranging between zero and infinity, according to the relative positions of the two groups of wave-lengths in the spectrum.

Another point of importance is that, for certain standard and fairly representative conditions, the function—visibility curve—connecting radiant energy for equal photometric brightness, with wave-length has been accurately determined for the average or normal eye. This being the case, it is possible, for these conditions, to deduce one measure from the other; and the eye can be regarded as a selective radiometer having a known calibration curve. Conversely,

Ives and Kingsbury[1] have carried out extensive experiments on a so-called 'physical photometer' or 'artificial eye' in which an absorbing solution, or equivalent arrangement, having a selective transmission corresponding closely to the normal visibility curve, is placed between the light source and a thermopile. The results obtained, although radiometrically determined, are actually proportional to the photometric brightness of the stimuli measured. The visibility (or stimulus) coefficient, which is plotted—as a function of wave-length—in Fig. 2, is directly proportional to the quotient obtained by dividing the photometric measure by the corresponding radiometric measure.

The average visibility function may now be considered a reliable technical asset of the investigator in vision.[2] Strictly

[1] Ives, H. E., and Kingsbury, E. F., 'Physical Photometry with a Thermopile Artificial Eye,' *Phys. Rev.*, 1915 (2), 6, 319–334. Also, Ives, H. E., 'A Precision Artificial Eye,' *ibid.*, pp. 334–346.

[2] Average 'visibility curves' for normal cone vision and flicker photometry have been determined by Ives, Nutting, Bender and others. Ives employed 18 observers and later 25 more. Nutting used 21 in all. The relative visibilities (brightness/relative energy) of various parts of the spectrum, as found by these two investigators, are given in the following table:

Wave-Length $\mu\mu$	400	410	420	430	440	450	460	470	480	490
(Nutting)	.002	.003	.008	.012	.023	.038	.066	.105	.157	.227
(Ives)	—	—	—	—	.029	.047	.073	.107	.154	.235
Wave-Length	500	510	520	530	540	550	560	570	580	590
(Nutting)	.330	.477	.671	.835	.944	.995	.993	.944	.851	.735
(Ives)	.363	.596	.794	.912	.977	1.000	.990	.948	.875	.763
Wave-Length	600	610	620	630	640	650	660	670	680	690
(Nutting)	.605	.468	.342	.247	.151	.094	.051	.028	.012	.007
(Ives)	.635	.509	.387	.272	.175	.104	.068	.044	.026	—

Wave-Length	700	710	720	730	740	750	760
(Nutting)	.002						
(Hyde and Forsythe)	.00282	.00137	.00068	.00033	.00017	.00008	.00003
Wave-Length	770						
(Hyde and Forsythe)	.000015						

Nutting's results are plotted in Fig. 2.

The values from 700 to 770 $\mu\mu$ are calculated from Hyde, E. P., and Forsythe, W. E., 'The Visibility of Radiation in the Red End of the Visible Spectrum,' *Astrophys. Journ.*, 1915, 42, 285–294. Nutting's results appear in his paper, 'The Visibility of Radiation,' *Trans. Illum. Eng. Soc.*, 1914, 9, 633–643; also in the *Phil. Mag.*, 1915,

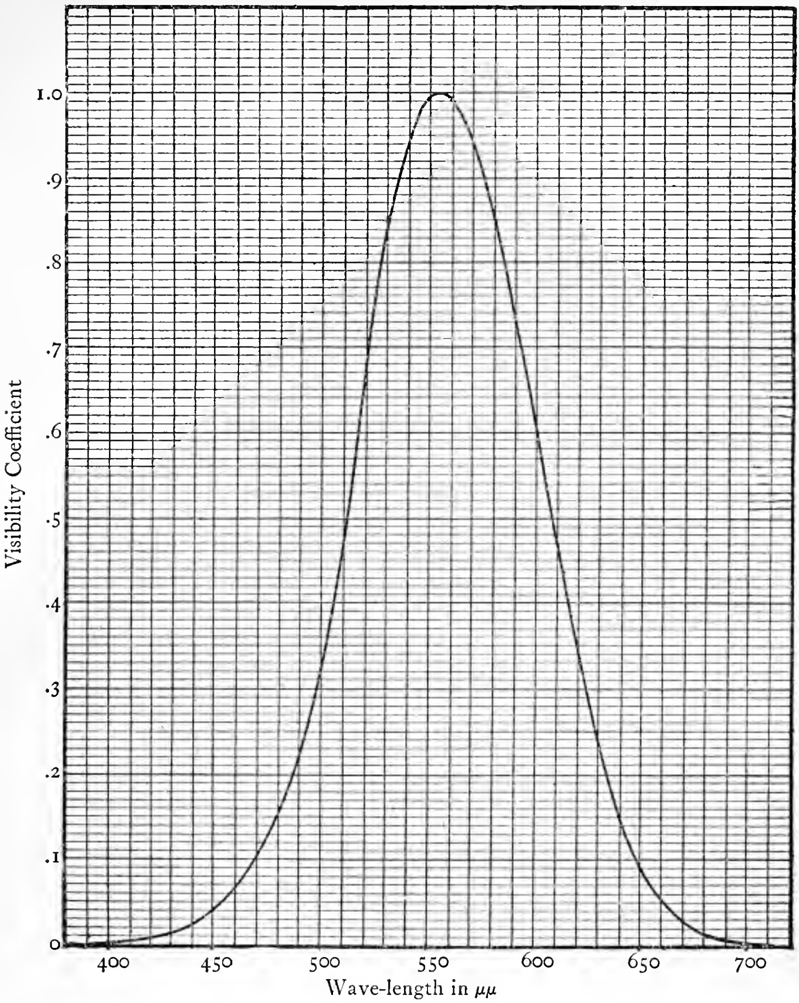


FIG. 2. The Visibility Curve According to Nutting. This curve is an accurate plot of the visibility values given below, representing the average results obtained from 21 subjects. If the photometric intensity of any 'homogeneous' visual stimulus is divided by the appropriate visibility coefficient (read off from the curve), and multiplied by the 'mechanical equivalent of light' (.00132 watts per lumen), the resulting value will be the radiometric intensity of the stimulus. By the converse procedure radiometric intensities can be changed into photometric values. Nutting's data were obtained by flicker photometry, and at a standard intensity of about 75 photons (see text).

speaking there will be a special visibility curve for each individual, for each species, for each absolute intensity of illumination of the retina, for each position on the retina, for each state of adaptation (whether general or specific) and for each method of photometry. However, comparison of the curves for different conditions does not, in general, reveal 'order of magnitude' discrepancies. It is very probable that the modal visibility curve for a given visual type is

(6) 29, 301. See also: Ives, H. E., 'The Spectral Luminosity Curve of the Average Eye,' *Phil. Mag.*, 1912 (6), 24, 853–864; and Bender, H., 'Untersuchung am Lummer-Pringsheimschen Spektralflickerphotometer,' *Ann. d. Phys.*, 1914 (4), 45, 105–144.

According to Nutting, the equation of the visibility curve is

$$V = V_m R^a e^{a(1-R)} = V_m v_\lambda, \tag{1}$$

where $R = \lambda_{max}/\lambda$, λ_{max} being the wave-length for the maximum ordinate of the curve, approximately 555, and λ the wave-length for which the visibility is to be calculated. $a = 181$; V_m is the visibility at the maximum and e is the base of the natural system of logarithms. V_m is the ratio of the lumen to the watt (unit of power) at the wave-length having maximum visibility, the reciprocal of the so-called mechanical equivalent of light, for this wave-length. Since the problem of making an accurate determination of the mechanical equivalent of light has only recently been attacked, there is still considerable lack of agreement between authorities, as to its value. Nutting (*loc. cit.*) finds .00120 watts per lumen; Ives ('Luminous Efficiency,' *Electrical World*, 1911, 57, 1565–1568), calculating certain data of Nernst's, gets .00125. Ives, Coblentz. and Kingsbury ('The Mechanical Equivalent of Light,' *Phys. Rev.*, 1915, (2), 5, 269–294) give .00159; Langmuir ('The Characteristics of Tungsten Filaments as Functions of Temperature,' *Phys. Rev.*, 1916 (2), 6, 302–330), .00121; and Hyde, Cady, and Forsythe ('The Candle Power of the Black Body and the Mechanical Equivalent of Light,' *Jour. of the Franklin Inst.*, 1916, 181, 420–421), .00132. The last value, which was calculated on the basis of Nutting's visibility data, was obtained from especially satisfactory data.

The value of the expression, $R^a e^{a(1-R)}$, in Nutting's equation should be equal, for a given wave-length, to the corresponding value in the table, *i. e.*, it gives the relative visibility, or visibility coefficient, of the selected wave-length, the visibility at the maximum being taken as unity. When multiplied by V_m (757), the values in the table, or those obtained from the curve in Fig. 2, or from the equation, will yield the visibility in absolute terms, *i. e.*, in lumens per watt.

The lumen is the unit of luminous flux, and the candle-power of any light source in a given direction is equal to the number of lumens emitted by it per unit solid angle, in this direction. In visual psycho-physiology, however, we are concerned immediately, only with units of brightness, which involve the distribution of candle-power over a surface. Brightness is expressed in lumens per unit solid angle per unit area of radiating surface. Accordingly, the reduction of visual stimulus values from photometric to radiometric terms, means the conversion of *candles per square meter*, to *watts per steradian, per square meter*. This latter quantity will be proportional, other things equal, to the energy impinging upon the retina, per unit time, per unit area.

practically independent of everything except the degree of participation, in the response, of rod process as compared with cone process. The problem of separating the effects contributed by these two distinct visual mechanisms, under all conditions, is one upon which further important advances are still possible. The standard visibility values now refer properly only to the normal, trichromatic, cone process.

Ultimately the technical definition of photometric brightness will probably be such that the brightness of any visual stimulus differing in quality from the standard will be determined only by its radiant power, and by the standard 'average visibility' of the given radiation. This standard specific visibility value will be determined by finding the mean values for a large number of normal human individuals under conditions most favorable for photometric comparison, and representing the most common circumstances of vision. On such a basis, with tested normal observers working[1] under the

To accomplish this reduction, the photometric value of the stimulus, expressed in candles per square meter, must be divided by the absolute visibility of the radiation constituting the stimulus. Thus, if b is the brightness measure in question, and v_λ is the proper visibility coefficient (obtained from the table or from the plot), we have, for watts/(steradian × square meter):

$$(2) \qquad w = \frac{b}{V_m v_\lambda} = L\frac{b}{v_\lambda} = .00132\, b/v_\lambda,$$

where L is the mechanical equivalent of light. The same quantity, expressed in *ergs*/(second × steradian × square meter), would be

$$(3) \qquad w' = 1.32 \times 10^4\, b/v_\lambda,$$

since one watt is equal to 10^7 ergs per second. The radiometric intensity of the radiation striking the retina can be calculated from w or w', by multiplying either of these quantities by a factor, the magnitude of which is determined by the size of the pupil, the focal length of the eye, and other ocular variables (*vide infra*).

[1] In order to select observers whose color vision approximates the established averages, H. E. Ives has devised color filters intended to provide a simple and effective test of the normality of any individual. See his papers, 'On the Choice of a Group of Observers for Heterochromatic Measurements,' *Trans. Illum. Eng. Soc.*, 1915, 10, 203–209; 'Experiments with Colored Absorbing Solutions for Use in Heterochromatic Photometry,' *ibid.*, 1914, 9, 795–814; 'Additional Experiments on Colored Absorbing Solutions for Use in Heterochromatic Photometry,' *ibid.*, 1915, 10, 253–259; 'A Method of Correcting Abnormal Color Vision and Its Application to Flicker Photometry,' *ibid.*, 1915, 10, 259–271. The Ives-Kingsbury test filters were employed by Crittenden and Richtmyer in their recent extensive investigations of flicker photometry.

specific standard conditions visual photometry can be construed as a special, convenient method of radiometry.

Ferree and Rand[1] claim that for the purposes of scientific physiological optics, the intensity of visual stimuli should be measured, and equated, in radiometric and not generally in photometric terms. If this is taken to mean that, at least with present methods, the ultimate basis of standardization of intensity measures is radiometric, the claim may be admitted. If, on the other hand, it is meant that direct radiometry is required in the psychological laboratory, for the immediate control of visual stimulus intensities, and that photometric measurements and equations should be rejected, we must certainly dispute the contention.

It would even be possible to argue that the radiometric treatment of the intensity of the stimulus is not only technically difficult, and perhaps lacking in immediate interest for the majority of problems, but is arbitrary. To measure radiation in terms of energy is to determine how much heat it can generate. Why not consider radiation in its immediately interesting context, and try to discover how much luminosity, or visual brightness, it can generate? In other words, let us replace the thermopile by a retina, and find the law of distribution of brightness in the spectrum of the stimulus which we are using. Then we shall be able to select visual stimuli of different wave-length, but of equal brightness.

From a strictly physical point of view the argument that the standardization of visual stimuli in terms of the temperature effect of the radiation is of no more fundamental significance than its standardization in terms of brightness, cannot be considered valid. In the first place, if we measure radiation intensity in terms of energy, and employ ideal instruments, our results will be independent of the instruments themselves; every radiocalorimeter which will absorb

[1] See Rand, G., 'The Factors that Influence the Sensitivity of the Retina to Color, etc.,' *Psychol. Mongr.*, No. 62, 32–40, 1913. Also: Ferree, C. E., and Rand, G., 'A Note on the Determination of the Retina's Sensitivity to Colored Light in Terms of Radiometric Units,' *Amer. J. of Psychol.*, 23, 328–333.

all of the radiation and convert it all into heat will give us the same measure, and the measure can never be greater than this. In the case of the retina and brightness, however, everything depends on the instruments, and there is, so far as we are aware, no maximally efficient instrument. In general, the energy of a physiological response is not derived from the stimulus, but is merely released by it. Only an almost infinitesimal fraction of the total gamut of electro-magnetic waves is capable of producing any brightness at all. There is nothing inherently distinctive about this fraction, either, so that, strictly speaking, its brightness, or its luminosity, is a property of the eye and its appendages rather than of the radiation itself.

Moreover, a measure of brightness cannot be regarded as being in absolute units unless it is referred to a definite radiant power (by use of the standard visibility values). It may be possible, ultimately, to measure the intensity of physiological—and even psychological—response in absolute units, but there is no established technique for doing this at present, so far as the results of photometry are concerned. If it were certain, or even likely, that this response intensity was the same for all individuals for a given intensity of the photometric standard, the response to the standard could be adopted as an absolute norm. As matters stand it can only be regarded as a relative norm.

In spite of all these objections, however, it is the writer's opinion that, in general, the photometric equation of intensities is preferable to the radiometric. This opinion has both a theoretical and a practical basis. The latter consists primarily in the extreme difficulty of making reliable radiometric measurements, as compared with the ease with which photometric equations are established, especially when the flicker method (*vide infra*) is employed. Radiometric and photometric measurements have, in general, about the same precision, somewhat better than one per cent. However, radiometers measure indiscriminately the energy of all types of radiation, whether visible or not, so that unless great care is taken to eliminate infra-red and ultra-violet rays, the

results are apt to be wholly misleading, considered as estimations of stimulus intensity, since these special rays are not visual stimuli.[1] Closely related with this is the fact that stimuli equated in energy over the whole spectrum will have a brightness so disproportionate as, in many cases, to make their visual comparison almost impossible. For example, under these conditions, a stimulus of wave-length 550 $\mu\mu$ would be about ten thousand times brighter than one of wave-length 750 $\mu\mu$. One of these stimuli might be just at the threshold, and the other would then be dazzlingly bright.

The purely practical difficulties of radiometry can be conquered by a careful technique. However, considering the interconvertibility of photometric and radiometric measurements, when the former are made by normal observers, it would seem unnecessary for the psychologist to trouble himself with these delicate procedures. *Moreover, whether the observers are normal or abnormal, there are certain theoretical advantages possessed by photometric equations, which cannot be ascribed to radiometric equations.*

Visual response consists of a series of stages or phases following each other in time, each phase having a number of more or less independent dimensions, the exact values in which, however, are determined primarily by the values of corresponding dimensions in the preceding stage. These stages, in their temporal order are, roughly: (1) visual object, (2) radiation from the object entering the eye, (3) the retinal image, (4) the receptor process, (5) the neural stimulation, (6) the afferent impulse, (7) the adjustor (or central) process, (8) the efferent impulse, (9) the effector process. In addition to these, and a function of some, and perhaps all, of the stages, there is: (*a*) visual experience. The values of the variables in each stage at any moment may be regarded as mathematical functions of the values of the variables in preceding stages at earlier moments.

[1] This statement may perhaps be questioned as applied rigorously to the ultra-violet, but it is the infra-red rays which are most bothersome in the radiometric measurements of visible radiation.

Now for the theoretical analysis of the total process of vision, which must be regarded as the ultimate scientific aim, the problem of the interrelation of what may be called the internal *factors* of the response (succeeding and including the receptor process), is probably of more importance, and is certainly far more difficult, than that of determining the relation between the stimulus variables (1 to 3) and certain internal factors. To separate the effects of the various variables in any internal stage of the response, certain of these variables, directly, should be held constant while others are subjected to change. It must be true that any two stimuli which produce the same effect upon a retinal receptor will have an equal value for all succeeding stages of the response. This will probably hold approximately, also, when the similarity of effect applies to only one dimension of the receptor process, if the subsequent influence of this dimension is considered in isolation from that of the others. The known facts provide us with some guarantee that photometric equations do establish such an equivalency—approximately—for the general dimension of intensity of the response, whereas there is no doubt whatsoever that stimulus energy equations fail in this respect.[1]

Such photometric equations would be relatively independent of specific, individual, and momentary variations in the sensitivity of the visual system, and hence would correct automatically for these variations, provided, of course, that the equations are made by the subject for whom the measured lights are later to be employed as further stimuli. Among themselves a system of intensity measures based upon photometry by *individuals selected at random*, must of course be considered as constructed with reference to any one subject's absolute sensibility to the standard of luminous intensity, as a norm.

However, when photometric equations are made by a

[1] The exact physiological nature of response intensities, apart from experienced *luminosity*—or other sensory attributes—remains, of course, somewhat vague. It may be a question of equality of reaction velocities, equality of energy released or absorbed per second, or—what is more likely—the concentration of certain ions in the respector cells.

tested, normal observer,[1] they refer back automatically to radiation intensities as a basis, since photometric values obtained under standard conditions must be equal to the corresponding radiometric values, multiplied by the appropriate visibility factor. Consequently, the radiant power of stimuli of wave-length constitution different from that of the standard light can be obtained by dividing the photometric quantity by the factor in question. If a normal observer is not to be found, methods can be applied for correcting the measurements so that they will coincide approximately with the normal.[2]

It is perfectly obvious, of course, that certain problems—such as those of visibility—involving the relationship between the receptor process and the stimulus, must be settled on the basis of direct energy measurements. Our contention here—in summary—is merely that such problems form a relatively small part of the whole group of questions faced by the psycho-physiologist in vision, and that for the remaining group, energy equations will yield results difficult, if not impossible, of interpretation.[3] Moreover, the existence at present of reliable determinations of visibility and of standard methods for applying them to photometric results makes radiometric and photometric measurements largely inter-convertible. The simple technique of photometry recommends it to the psychologist, and at the same time permits him to express the conditions of his observations in such a manner that they can be interpreted and reproduced by others.[4]

[1] The appropriate tests have been described by Ives. See the references above.

[2] See Ives, *Trans. Illum. Eng. Soc.*, 1915, 10, 259–271. Also Crittenden, E. C., and Richtmyer, F. K., 'An "Average Eye" for Heterochromatic Photometry, and a Comparison of a Flicker and an Equality-of-Brightness Photometer. *Trans. Illum. Eng. Soc.*, 1916, 11, 331–372.

[3] In the writer's estimation, the questions raised primarily by Ferree and Rand, *viz.*, the peripheral limits of color sensitivity in the retina, fall mainly in the group where energy equations will complicate, rather than simplify the investigation. *Cf.* Baird, J. W., 'The Phenomena of Indirect Color Vision,' *Psychol. Review*, 1914, 21, 70–79.

[4] It should be borne in mind that ordinary direct radiometry, as carried out by means of a thermopile or bolometer, yields relative energy measures, only. The technique of absolute radiometric measurements, or radiocalorimetry, is even more difficult than that of relative methods.

III. The Problem of Heterochromatic Photometry, and the Justification of the Flicker Photometer

If the procedure of photometry involved in general only the equation of stimuli having the same energy distribution curve in the spectrum and hence the same color, the problem would be a very simple one. As a matter of fact, it is very seldom that two visual stimuli to be equated have the same distribution curve, and this, of course, is never the case when the main conditions which we have been considering above are in question, for under such conditions radiometry and photometry are *equivalent* methods. Moreover, photometry in general is color photometry, since none of our light sources are strictly 'white' and even sunlight varies in hue.

The problem of the proper method for the equation, in respect to brightness, of two lights of different color, has been under discussion for many years, but has never received more attention than is being devoted to it at the present day. This problem is partly experimental and partly logical, and more progress has been made on the experimental than on the logical side of the question. The latter aspect of the problem concerns primarily the definition of 'equal brightness,' and the establishment of a scientific criterion for choosing a satisfactory method of heterochromatic photometry.[1]

The term 'brightness' may now be regarded as having been definitely appropriated by the photometrician to designate that aspect of illumination questions which has immediate reference to the effect of a given stimulus on a given eye. The brightness of an illuminated surface depends upon a point of view. When lights of different colors are compared the brightness must also depend upon the visibility curve of the eye, and this visibility curve will not, in general, be the same for different methods of heterochromatic photometry. *It would seem advisable to employ the term brightness to express*

[1] The criteria have been considered by Ives, *Phil. Mag.*, 1912 (6), 24, 153–157. An excellent discussion of the problem of heterochromatic photometry will be found in M. Luckiesh, 'Color and Its Applications,' 1915, Chap. IX. This book contains much useful data for the investigator of vision.

the photometric value obtained for any visual stimulus by the standard method of photometry. For example, if we adopt the flicker method as our standard procedure the flicker value of a visual stimulus will be its brightness.[1]

Ordinary or direct comparison photometry is based upon an equation of luminosities. The criterion of 'equal luminosity' is not ambiguous when the lights to be compared are of the same color (hue and saturation) since in this case the task of the photometrician is merely to find two amounts of radiation which produce entirely similar sensations. However, when the two radiations are such as to condition a noticeable difference in color, it is necessary to discriminate inspectively between the experiential dimensions of hue and saturation, on the one hand, and that of *luminosity* on the other.

Now there seems to be a general psychological law that the distinctness of any experiential (or qualitative) dimension changes in parallel with the degree of similarity of two compared experiences in all other dimensions. For example, if luminosity, hue and saturation are three dimensions of a visual sensation (*per se*) the threshold for the perception of a difference in luminosity between two sensations will be greater the greater the concomitant differences in hue or saturation, or both. This principle may depend upon a weakening of our powers of discrimination or it may indicate that the dimensions, as such, are to a certain degree mutually dependent and unreliable. In other words, the meaning of the term 'equal luminosity' may become *ambiguous* in proportion as two compared sensations differ in color.[2]

[1] Since the flicker value and direct comparison value of a given colored light do not, in general, agree, this definition would mean, on the basis of flicker photometry, that equally bright lights do not always generate equal luminosities. This usage of 'brightness' of course conflicts with some traditional phrases, such as 'equality of brightness,' but is in line with the modern development of photometric nomenclature.

[2] This conception may perhaps be expressed by saying that a system of ideal Cartesian axes for the determination of values in these dimensions would not be a set of mutually perpendicular *lines* intersecting in a single point, but would consist of a group of (perhaps truncated) *cones*, with their axes perpendicular and with their apices meeting at the origin of coördinates. This would mean that all differences in luminosity, hue, saturation, tint, etc., are to a certain extent indeterminate, and that

The hypothesis that in heterochromatic comparisons of luminosity it is not essentially the process of judgment but rather the conception of the dimension of luminosity itself which is uncertain, surely represents the psychological facts of the case very well. Relations exist which can be judged categorically as *differences* in luminosity, but it is never possible to make a categorical judgment of *equality* of luminosity. The truthful judgment is always: "I cannot tell whether they are equal in luminosity or not." Langfeld[1] and others have shown that the results obtained by heterochromatic comparison depend radically upon the 'attitude of the observer,' or upon the 'criterion' for equality of luminosity which is adopted. This 'criterion' would amount, on our theory, to a redefinition of the term 'luminosity' for the special comparison involved.

The existence of this uncertainty as to the meaning of the term 'equal luminosity' for lights of different color should lead one to conclude that equality of luminosity, and consequently the method of direct photometric comparison, cannot be regarded as furnishing a satisfactory test of equality of brightness for such lights. The uncertainty of course makes itself evident objectively in the lack of precision of photometric measurements made by this method, and in the lack of agreement between different normal observers.

this indetermination becomes greater in any given dimension the greater the established differences in other dimensions. The facts are such that this hypothesis involves us in fewer serious assumptions than one which refers the difficulty to a fallibility of the discriminative function.

If we adopt the hypothesis in question, it follows that the term 'equal luminosity' necessarily becomes more and more ambiguous the greater the difference in hue and saturation, between two compared sensations. In other words, 'equal luminosity' cannot represent any definite condition of affairs, except the absence of 'unequal luminosity,' a requirement which could be satisfied in many different ways. On this basis, since the photometric equation of brightnesses depends upon the equation of sensory luminosities, the ordinary procedure of 'direct comparison,' or 'equality of brightness,' photometry would tell us unequivocally when two brightnesses were unequal, but not unequivocally when they were equal.

Recent experiments by the writer show that with large color differences between compared visual fields, the just noticeable difference in brightness may exceed 20 per cent. See 'The Heterochromatic Brightness Discrimination Threshold,' *Journal of the Franklin Inst.*, 1916, 182, 112–115.

[1] Langfeld, J. S., 'Ueber die Heterochrome Helligkeitsvergleichung,' *Zeitschr. f. Psychol.*, 1909, 53, 113–179.

If we reject the method of direct comparison we must cast about for another procedure which measures approximately the same quantities, but which depends upon more definite psychological criteria, and possesses a better objective precision. A large number of careful experiments, due primarily to H. E. Ives,[1] have proved unquestionably that *flicker photometry* satisfies this requirement. It has been shown by Ives that (1) "the flicker method is more sensitive than the equality of brightness method, where different coloured lights are compared," and that (2) "the results by the flicker method are more reproducible than those by the equality of brightness."[2] Crittenden and Richtmyer, in recent work, find the mean variation of the results of 114 observers by the 'equality of brightness' method to be 1.9 per cent. and by the flicker method .6 per cent. for the same stimuli. They say: "With regard to certainty of measurement the flicker photometer shows a decided advantage even with small color differences. With more experienced observers, specially selected, this advantage would probably be materially reduced, but would not be entirely lost, because even when an observer makes consistent settings on the equality photometer the relation of his settings to those of the normal observer is uncertain. . . . with the flicker any observer of fair ability can make definite sets even with large color differences whereas on the Lummer-Brodhun ['equality of brightness'] photometer it is only the exceptional observer who can do so. . . . The flicker photometer affords a means of relatively precise comparison between lights of all degrees of color difference, and makes possible the use of test readings for which average values, which should be highly reproducible, can be established."

Ferree and Rand[4] have attacked the flicker method on

[1] Ives, H. E., 'Studies in the Photometry of Lights of Different Colours,' *Phil. Mag.*, 1912 (6), 24, 149–189, 352–370, 744–751, 845–864. See also the recent elaborate investigations, with 114 subjects by Crittenden, E. C., and Richtmyer, F. K., 'An "Average Eye" for Heterochromatic Photometry, and a Comparison of a Flicker and an Equality-of-Brightness Photometer.' *Trans. Illum. Eng. Soc.*, 1916, 11, 331–367.

[2] *Loc. cit.*, p. 177.

[3] Ferree, C. E., and Rand, G., 'A Preliminary Study of the Deficiencies of the Method of Flicker for the Photometry of Lights of Different Colors,' *Psychol. Rev.*, 1915, 32, 110–163.

the ground that the results which it yields disagree with those of the 'equality of brightness' procedure. They explain the discrepancy in terms of the different rates of growth and decay of sensation for the different colors. This criticism of the use of the flicker method—and the considerations on which it is based are not new[1]—presupposes that the object of a method of photometry is to measure light in terms of equated luminosities, and that to justify the flicker procedure, it is necessary to prove that the *flicker value* of a light agrees within the limits of precision of the measurements with its value as determined by direct comparison. It is one thesis of the present article that this presupposition is arbitrary and scientifically questionable.

In the first place, it must be admitted that an acceptable method of photometry must yield results which agree *approximately* with those of the 'equality of brightness' procedure; the disagreement should certainly at no point be greater than a single order of magnitude. The reason for this requirement is to be found in our conviction that luminosity is closely proportional to the intensity of the response, and to the utility of the radiation, the variables in which we are fundamentally interested. However to require accurate agreement between the two sets of results would be unreasonable, first, because of the relative ambiguity of 'equality of luminosity,' and secondly, because it is improbable that luminosity or any other property of the response which we may select as a basis for establishing a photometric balance, can be considered immediately indicative of all phases of the intensity of the response, without correction for the special conditions of its utilization.

Ferree and Rand do not claim to be pioneers in the proof that the results of the flicker method and that of direct comparison do not agree. Ives[2] found that the curves showing the distribution of brightness in the spectrum, as determined by the two methods, differed in a number of very

[1] See Luckiesh, M., 'On the Growth and Decay of Color Sensations in Flicker Photometry,' *Phys. Rev.*, 1914 (2), 4, 4–6.

[2] *Loc. cit.*, pp. 177–178.

definite ways. However, in general these differences were small. Crittenden and Richtmyer say: "In regard to relative results there appears to be no room for doubt that for sources having relatively high intensity at the blue end of the spectrum the values given by the flicker photometer as here used depart appreciably from those obtained with the Lummer-Brodhun as used in common practise, the difference being of the order of 2 per cent. at the higher efficiencies reached by the present gas-filled lamps."[2]

Considering the probable complexity of the processes of growth and decay of color sensation it would indeed be surprising if the results obtained by the flicker photometer depended directly, without correction, upon the luminosity value of a light. Very interesting theoretical studies on this question have been made by Ives and Kingsbury.[2] Successive contrast must also introduce complications,[3] just as simultaneous contrast must influence the results of direct com-

[2] The same authors continue: "It is, however, hardly proper to assume that the results obtained by either photometer are 'right' and anything different is 'wrong'; the equality-of-brightness method of measurement is undoubtedly more closely related to the way in which the light is used, but it is by no means established that the method correctly indicates the relative usefulness of two kinds of light. It must be recognized that there is no one definite 'correct' ratio between the intensities of two lights of different color. . . . The specification of conditions of measurements must be more or less arbitrary, and the results obtained cannot be expected to be an exact indication of the value of different kinds of light under different conditions. Before we shall know much about the relative usefulness of different kinds of light much more experimental work must be done; an important requisite for such investigations or any others involving the comparison of the intensity of lights by very different color is a method which will enable different experimenters to make consistent measurements of the quantity which must serve as a basis for the comparison of their results. The usual equality-of-brightness method of comparison certainly does not fulfill this requirement; the flicker photometer at present is the most promising method available." "Comparison of actual tests made in the routine work of the laboratory shows that even with relatively small color differences a given accuracy of reproduction of results requires several times as many measurements with the equality-of-brightness or the contrast photometer as with the flicker; moreover the tests considered were made by observers who had had much experience with the contrast photometer and very little with the flicker."

[2] Ives, H. E., and Kingsbury, E. F., 'The Theory of the Flicker Photometer,' *Phil. Mag.*, 1914 (6), 28, 708–728. This article represents clearly the correct way to attack visual problems, and contains concepts of fundamental importance for a large number of visual phenomena.

[3] Cf. Luckiesh, *lec. cit.* p. 6–8.

parison.[1] The theory of the flicker photometer as developed by Ives and Kingsbury, indicates that its results should approach asymptotically those of the direct comparison method as the intensity of stimulation increases. This deduction is in harmony with fact.

The essential point to be established, however, is this: the one necessary requirement of a method of measurement is that it shall permit the accurate reproduction of experimental conditions. As shown by the work of Ives and other modern students of the problem of heterochromatic photometry, this requirement is met by the flicker method, and not by any other procedure which has been adequately tested.[2] As a consequence, we are forced at present to accept the flicker method as our standard procedure, and to define photometric brightness in terms of its results regardless of the fact that these results differ somewhat from those obtained by the criterion of equal luminosity.

It is possible that we shall ultimately find some reliable method for measuring directly what may be called the *true response intensity*, for a specified stage of the response. The exact functional connection between this quantity and flicker value can then be determined, so that the true intensity can be deduced from the flicker value. Possibly the nature and magnitude of this correction can be deduced from the theory of the flicker photometer. For many purposes, however, the conversion of true physiological intensity would prove useless, since a knowledge of the laws connecting other properties of stimuli, *e. g.*, such as acuity values, with flicker value should be as serviceable as a knowledge of the relation between these properties and true intensity. For theoretical purposes, of course, equations of true intensity are highly desirable, but it is by no means certain that the method of direct com-

[1] See Bell, L., 'Some Factors in Heterochromatic Photometry,' *Electrical World*, 1912, 59, 201–203.

[2] Ferree ('A New Method of Heterochromatic Photometry,' Journal of Exp. Psychol., 1916, 1, 1–13) has recently proposed a new procedure for which he claims remarkable accuracy. Since the method is a peculiar one, not immediately suggestive of reliability, it will have to be more carefully investigated before it can be taken seriously.

parison would not demand corrections, in this respect, as great as those required by flicker. Moreover the flicker results furnish a reliable measure upon which to base such a correction, which cannot be said of the results obtained by direct comparison.

As already pointed out, the general problem of photometry is that of establishing heterochromatic equations, since color differences are the rule rather than the exception, for the lights which need to be compared. However, when the color difference is less than some critical amount, the method of direct comparison becomes more sensitive than that of flicker. But, as color difference disappears, the results of the flicker procedure approach identity with those of direct comparison, so that for lights differing only slightly in color from the standard it is immaterial whether we define photometric brightness in terms of flicker or of equal luminosity values, and in such case—as in general—we will naturally employ the method which is the most reliable.

IV. The Influence of the Pupil on the Intensity of the Visual Stimulus, and the Definition of a Standard Unit for Specifying the Intensity of Visual Stimulation

In a previous article[1] the writer has emphasized the importance of the artificial pupil in the control of the intensity of visual stimuli. The actual visual stimulus is the retinal image, and the illumination which this image represents is always proportional to the area of the pupillary opening. The normal range of variation of pupillary area is from 1 to about 16, so that neglect to control the size of the pupil would introduce a factor of uncertainty into our measurements of the intensity of the visual stimulus—or our equations of response energies—perhaps as great as 1,600 per cent. This is not large compared with the range of external illuminations, but it is enormous compared with accuracy with which intensities can be determined by photometry.

[1] Troland, L. T., 'The Theory and Practise of the Artificial Pupil,' *Psychol. Rev.*, 1915, 22, 167–177.

All careful workers in vision have recognized this fact, and have taken pains to keep the pupil opening constant, and to state the pupillary diameter as one of the conditions of their experiments. This is true of Ives's studies on the flicker photometer, of Nutting's determination of the visibility function, and of the earlier careful measurements of König on visibility and difference threshold. There may be some problems in vision for which the *order of magnitude* of the intensity alone is of importance, but with entire neglect of the pupillary size we cannot even insure a knowledge of the exact order of magnitude of the retinal illumination. In the present state of visual science it is not safe to assume that a determination of the order of magnitude of the intensity, or an establishment of intensity equations to a first order of approximation, is generally adequate.

It is perhaps needless to say that these considerations apply with equal force both to radiometrically and photometrically determined intensities.

When the pupil size is known it is convenient to express the intensity conditions in terms of unit pupil area. This is done by both Ives and Nutting in the researches already referred to. In an extensive monograph on the laws of color adaptation, yet to be published, the present writer has expressed his intensity measures throughout in terms of a unit involving the pupillary area, and has proposed that this unit, called the *photon*, be adopted as the standard means of specifying the photometric intensity of visual stimulation conditions.

The definition of such a unit involves a number of independent considerations. The first of these concerns the mode of expressing what may be termed the *external intensity* of the stimulus. As a rule this external intensity is given, in descriptions of visual work, in terms of *meter-candles*. This is, of course, a mistake, since the meter-candle is a unit of *illumination*, *i. e.*, it measures the light *falling upon* a surface, whereas for purposes of visual experimentation, it is the light (density) *leaving* the surface *in the direction of the eye* which, alone, is of importance. To deduce this quantity from the illumination it is necessary to know the coefficient

of reflection, and the diffusing power of the surface. The quantity in question is called the *photometric brightness* of the surface in the given direction, and is measured in candles per unit area, or in *lamberts*. It is the only external photometric quantity which is of importance to the visual physiologist, since it is the brightness which determines the illumination of the retinal image.

It would be supererogatory to discuss this question of the measurement of the external intensity of the stimulus further in this paper, since it has been thoroughly treated in the recent and extremely useful article by Cobb, already referred to.[1] Suffice it to say that the first step is to determine this intensity in (let us say) candles per square meter, from the exact point of view to be employed by the subject in the experiment.

This determination should be made by a normal subject under standard conditions (or else corrected to the normal) and, in general, by the use of a flicker photometer. The photometer should be of the Whitman disk type, in which the alternation consists of an instantaneous substitution of the measured light for the standard at any one point of the retina, and in which the periods of presentation of the standard and the measured light are equal. Moreover, the speed of the photometer should be adjusted so as to give maximum sensibility, *i. e.*, should not be increased beyond the point at which color-flicker just disappears.[2]

The photometric equation should be established with the eye centered in front of an artificial pupil sufficiently small to prevent oscillations of the natural pupil from cutting off any of the light.[3] This pupil may be of any shape, provided its area is known, and in the standard case its axis should coincide with the line of sight. For stimuli produced by simple transmission or reflection arrangements, the circular pupil is preferable, while for spectral stimuli in which an

[1] *Psychol. Rev.*, 1916, 23, 71–89.

[2] The writer is at present making careful measurements of this critical speed for the spectral colors at various intensities, and with standard lights of various colors. See *Journal of the Franklin Inst.*, 1916, 181, 553–555.

[3] See the writer's article on the artificial pupil.

image of a slit is thrown upon the eye, a square, or otherwise rectangular, opening is more convenient and reliable. In the case of a slit image it is not necessary—in determining the photon value—that the image should fill the pupil, provided the light from the standard fills it. To establish a photometric equation the two retinal illuminations must always have the same brightness, and any increase in the effective pupil area for one will thus automatically be compensated for by an inversely proportional decrease in the external intensity needed for the equation.

Let us suppose that the measured brightness is b candles per square meter, and that the area of the pupil is p square millimeters. Then if r is the illumination of the corresponding retinal image (in meter-candles), we have

$$r = jpb, \tag{4}$$

where j is a factor depending upon the reflection, absorption and scattering of light in the eye, upon the angle of incidence of the rays with reference to the line of sight, and upon other influences to be considered below.

Suppose, now, that in an ideal case both p and b are equal, separately, to unity. Obviously in this case the value of the product, pb, will also be unity. For a given value of j this condition represents a definite retinal illumination, which will be duplicated whenever pb has the value of unity, whether or not the individual components have this value. Let us arbitrarily select this convenient intensity, for certain standard conditions, to be specified more in detail below, as the unit of physiological stimulus intensity, which may be known as the *photon*.

It should be clear from the above that if the intensity conditions of stimulation in visual experiments are expressed in photons they are unequivocally determined and can be directly compared, regardless of the actual pupillary size which was used, or the actual external brightness. To obtain this photon measure approximately it is only necessary to multiply the photometrically ascertained brightness—in candles per square meter—by the area of the pupil—calculated in square millimeters.

In order to determine accurately the physiological intensity of a stimulus surface under any desired condition, it is necessary to correct for variations in the factor j. The value of this factor may be regarded as summing up the

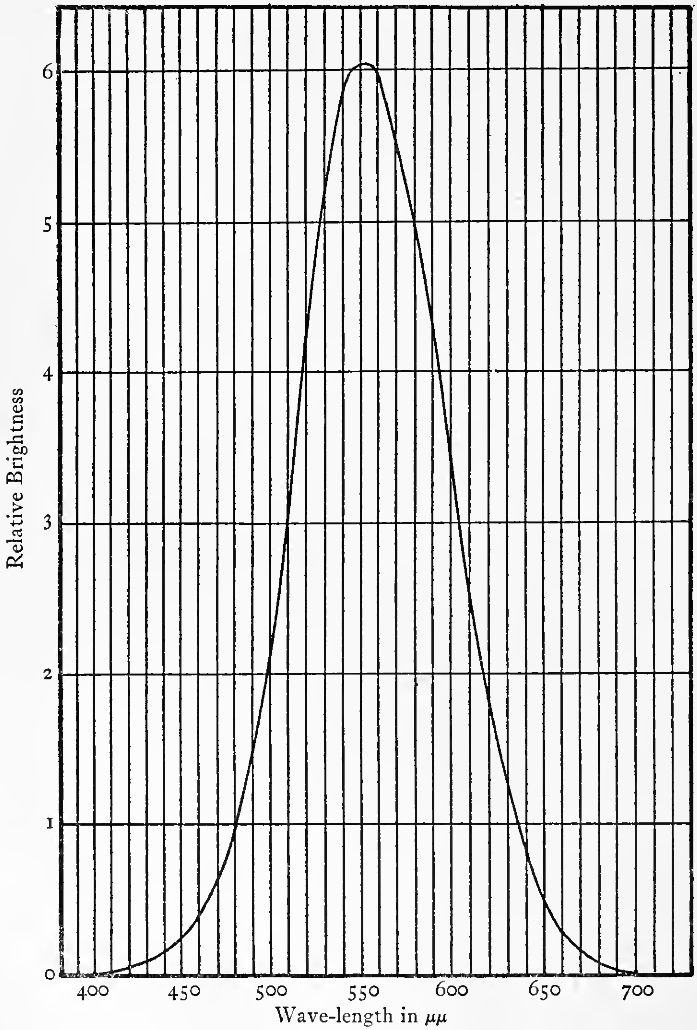


Fig. 3. Brightness Distribution Curve of the Light from a "Black Body" at 6200° Absolute. This curve was obtained theoretically by multiplying each ordinate of Fig. 1 by the corresponding ordinate of Fig. 2. It will be seen that it is practically identical in form with the visibility curve.

influence upon retinal illumination of variables other than the external brightness and the area of the pupil. The most important of these variables are the angle made by the direction of the stimulus surface with the line of sight, and the distance of the artificial pupil from the nodal point of the eye.

If ϕ is the angle in question (see Fig. 4) and the plane of the pupil is perpendicular to the line of sight, j must contain the factor: $\cos \phi$. On the assumption that intensity differences less than one per cent. may be neglected, the influence

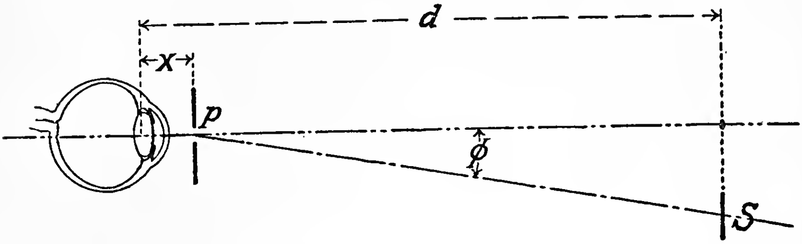


Fig. 4. The Influence of the Position of a Stimulus Surface and Artificial Pupil on the Intensity of the Retinal Image. See the text.

of angle can be disregarded, unless ϕ is greater than 8°, since the cosine of 8.1° is .9900. In the paper already referred to, the writer has estimated that the limiting angle for the use of the ordinary artificial pupil is about 53°. The cosine of 53° is .6018, the reduction in the physiological intensity of a stimulus viewed at this angle therefore amounting to about 40 per cent. of its value for direct vision.

For larger angles than this it is still possible to use the principle of the artificial pupil by means of an optical train which forms a sharp image of a small diaphragm within the natural pupil. This device has been employed by Cobb,[1] although not for peripheral stimulation. In an arrangement of this sort the area of the effective pupil depends entirely upon that of the diaphragm, the image of which falls within the natural pupil, but may be considered equivalent to the

[1] Cobb, P. W., 'The Influence of Illumination of the Eye on Visual Acuity,' *Amer. Journ. of Physiol.*, 1911, 29, 87.

pupillary area which was employed in photometering the light. In this case—provided, of course, the image clears the iris—there is no cosine effect.

It is possible to avoid the cosine correction in the use of the ordinary artificial pupil if the pupil is placed normal to the line passing through center of the stimulus surface and the nodal point of the eye, provided the surface in question is small. However, such an arrangement of the pupil presents no practical advantages, but only difficulties.

There is, of course, also a cosine effect when the natural pupil is employed, although it cannot be calculated diiectly from the external visual angle of the rays, on account of the refraction which occurs at the cornea. This tends to reduce slightly the average angle at which the rays strike the natural pupil, so that the reduction of intensity will not become appreciable at so small an external visual angle as with the artificial pupil.

In addition, for both natural and artificial pupils, there are other influences affecting the retinal illumination. The loss of light by reflection at the various refracting surfaces of the eyes increases with the angle at which the light impinges upon them. Besides this, the light strikes peripheral regions of the retina obliquely to the surface. Both of these effects reduce the peripheral retinal illumination. The peripheral regions, however, are reached by the light after passing through a somewhat thinner layer of absorbing material than is the case for the central regions, which would involve a slight relative increase in the illumination.[1] All of these factors would have to be taken into consideration in determining an accurate value for the physiological intensity of stimulation in the extreme periphery, although they can safely be neglected for fixated fields of 20° or smaller.

If the *photon* should be adopted as the universal unit for expressing the intensity of retinal stimulation it would of course be necessary to reduce the intensity to terms of this unit even when an artificial pupil is not used. This can be

[1] This is not by any means a complete catalogue of the factors influencing the intensity of the light which gets to the retinal receptors, but it will probably suffice for the present discussion of approximations.

accomplished, approximately, by multiplying the photometric brightness of the stimulus by the average apparent size of the natural pupil—the so-called *Eintrittspupille*.

Analysis shows that the illumination of the retinal image is not wholly independent of the distance between the artificial pupil and the eye. In general the illumination of the retinal image is nearly independent of the distance of the *object* from the eye.[2] This is a result of the fact that the area of the image changes in close proportion to the total light flux entering the pupil from the object. However, when the eye moves relatively to an artificial pupil, and the object is stationary with respect to the pupil, this compensation does not occur, since the total flux remains the same, while the area of the image alters.

If f is the focal length of the eye, d the distance of the stimulus surface from the nodal point of the eye, and x the distance between the artificial pupil and the nodal point of the eye, we can argue as follows. Take S as the area of the (small) stimulus surface, and b, as its brightness. Then the area of the retinal image will be Sf^2/d^2, and if p is the area of the artificial pupil, the total flux passing through it will be $Sbp/(d - x)^2$. Consequently, the illumination of the image must be proportional to

$$\frac{bpd^2}{f^2(d - x)^2}.$$

It is seen that if x is small compared with d its influence can be neglected.

Assuming that the effect of x upon the retinal illumination must be less than one per cent. to be negligible, we may solve the equation: $d^2/(d - x)^2 - 1 = 1/100$ to determine how small x must be made in order that its influence can safely be neglected. We find: $x = d/201.5$, or the distance between the nodal point of the eye and the artificial pupil must be less than 1/200th of the distance between the nodal point

[1] See Cobb, *Psychol. Rev.*, 1916, 23, 80–81. This proof does not hold accurately for objects close to the eye, since the area of the retinal image depends on the distance from the object to the nodal point of the eye, while the total flux of light contributing to the illumination of the image depends on its distance from a plane between the iris and the cornea, the nodal point of the 'reduced eye' lying posterior to the iris.

and the stimulus surface in order that its influence shall be negligible. Accordingly, it is necessary to adopt some standard position for the pupil. Optically, the natural position would be at the nodal point of the eye, since if the pupil were at this point, the retinal illumination would be independent of the distance of the stimulus surface. However, the natural pupil ordinarily lies about 2.7 mm. anterior to the nodal point and an artificial pupil can hardly be placed nearer than 10 mm. to it, or about 4 mm. from the cornea.

On account of the general necessity for correction—whatever the standard position adopted—it seems advisable to choose the plane of the nodal point of the eye, although no pupil ever does actually take this position. On this basis, the photon value is given with considerable accuracy by the equation:

$$i = \frac{pbd^2}{(d - x)^2} \cos \phi, \tag{5}$$

the significance of the variables being as already defined.

Of course, the above discussion must be considered only as approximative, but formula (5) will suffice for most purposes.

The following formal definition may be given of the *photon*, and of *the physiological intensity of a visual stimulus*. A *photon* is that intensity of illumination upon the retina of the eye which accompanies the direct fixation, with adequate accommodation, of a stimulus of small area, the photometric brightness of which, as determined by the standard flicker comparison and a normal subject, is one candle per square meter, when the area of the externally effective pupil, considered as lying in the nodal plane of the eye, is one square millimeter. The *physiological intensity of a visual stimulus* is its intensity expressed in photons. The photon is a unit of illumination, and hence has an absolute value in meter-candles.[1]

[1] The numerical magnitude of the photon, in meter-candles, and also its reduction to energy units will be considered by the writer in a further paper. It will obviously be subject to some variation from individual to individual.

V. Summary

The present paper is a somewhat discursive study of certain very general questions with regard to the measurement of the intensity of visual stimuli. The writer hopes that as a review of facts, as well as of problems, the paper will prove itself useful to the psychologist.

The various meanings of the term intensity are discussed, and the fundamental significance of photometric and radiometric measurements is considered, together with the relations which hold between radiant energy and light. Recent important empirical studies of these questions are summarized. On the basis of these facts and a theory as to the probable physiological significance of photometric equations, it is claimed that in general such equations will be more useful to the student of visual psycho-physiology than will radiometric equations.

The fundamental presuppositions of a method of photometry are then taken under consideration, and on the basis of recent very careful studies of the method of flicker, it is claimed that this method should be adopted, at least tentatively, as the standard photometric procedure, whenever two compared lights show a color difference.

A third aspect of the problem concerns the influence of pupillary size and other factors besides the external brightness of the stimulus, upon the illumination of the retinal image. In a preliminary discussion of this question, new considerations with reference to the use of the artificial pupil are presented, and the *photon*, defined as a unit of *physiological stimulus intensity*, is offered as a basis for the general standardization of conditions of visual experimentation, with regard to intensity.

This paper is written primarily for the experimental psychologist, rather than for the photometrician or illuminating engineer, but at the same time discusses fundamental problems of general interest.

March 29, 1916

A NEW METHOD FOR MEASURING REACTION-TIME

BY F. E. AUSTIN

Hanover, N. H.

The methods commonly adopted for measuring reaction-time, depend upon the use of an electromagnet operating some form of armature to which is attached a marker or stylus.

As is well known, every electromagnet has a reaction time of its own; that is, a certain interval of time elapses between the 'making' of the electrical connections closing the circuit, and the acting of the recording stylus. An interval of time also elapses between the 'breaking' of a circuit and the release action of an electromagnetic stylus arrangement.

Were the 'making' and the 'breaking' reaction times of a given electromagnetic stylus arrangement, of exactly the same duration, the apparatus could be relied upon to give consistent results. When, however, the lapse of time for the closing action is less than for the opening action an error is introduced in the record made by the apparatus.

Experiment seems to demonstrate that not only are electromagnet reaction times at make and break different, but that there is a considerable *variation* in the reaction time at make, for the same electromagnetic arrangement. This fact at once makes evident the introduction of a *variable* error in the results of many investigations not at all amenable to correction.

Were the variable error not a large one as compared with the quantities measured, its existence might be tolerated; the variable error has been found to be very large in proportion to such intervals as the shortest reaction times of individuals, and has led to the development of the new method designed to eliminate errors.

The method for measuring reaction time to be described is based upon the fact that the arc produced by an alternating

current between two electrodes is set up and extinguished twice during each cycle of alternations and also upon the accuracy with which the frequency of an alternating current may be indicated and observed by employing a so-called Frahm frequency meter.

If the image of an alternating arc be focused upon a photographic plate, and the plate be *exposed* while it is rapidly moved laterally in the focal plane, the impressions upon the plate when properly developed are as shown in Fig. 1; which

FIG. 1

shows two records made upon a moving film, by an arc produced between two ordinary electric-light carbons, by a so-called 60-cycle alternating current. It may be noted that the more rapid the motion of the film, the longer are the impressions or the dashes made by the light, and likewise the greater the distance between the light dashes.

A record produced by a 60-cycle alternating-current arc during one second would consist of 120 light dashes and a corresponding number of intervening spaces. Two dashes and two spaces would therefore occupy $\frac{1}{60}$ of the complete record for one second.

If a film should be moved at the rate of 60 inches per second, then 1 inch of record would correspond to $\frac{1}{60}$ of one second; which expressed as a decimal is $.016\frac{2}{3}$ second. By properly measuring the fractional part of one inch of such a record, less than $\frac{1}{60}$ of one second may be easily measured. For example $\frac{1}{4}$ inch of record as designated, would represent $\frac{1}{4}$ of $\frac{1}{60}$ or $\frac{1}{240}$ of one second; or expressed as a decimal $.004\frac{1}{6}$ second.

Should the film be moved at the rate of 120 inches or 10 feet per second, $\frac{1}{4}$ inch of record would represent $\frac{1}{480}$ of one second; or $\frac{1}{8}$ of an inch of record would represent $\frac{1}{960}$ of one second; or very nearly one one-thousandth of a second.

It is evident that the speed of the photographic film need not have a fixed constant value, but may vary considerably,

provided the *frequency* of the alternating arc *is known when any given record is produced* on the film. In other words the accuracy of record does not depend upon uniformity of motion of the surface upon which the record is made.

A frequency meter, such as the Frahm, consisting of carefully calibrated reeds is a very accurate indicator of frequency, and the essential requisite is to note the indication of the meter at the instant the sensitive film is exposed to the action of the light from the alternating arc.

To eliminate observational errors, a photograph (by flash-light if desirable) of the dial of the frequency meter can be taken *simultaneously* with the record on the moving film, but for ordinary investigations, a careful observer is sufficient.

Fig. 2 shows, in plan, the arrangement of apparatus used in obtaining reaction time for sight.

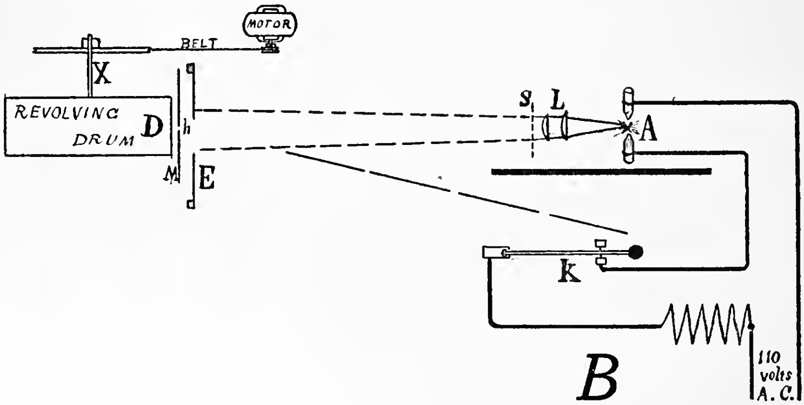


Fig. 2

The apparatus consists of an alternating-current arc A, a lens L to focus the beam of light from the arc upon the surface of a sensitive photographic film attached to the surface of a drum D that is arranged to revolve about an axis X. The drum is made of aluminum, is 6 inches wide and has a circumference of $45\frac{1}{2}$ inches; this means that a sensitive film $45\frac{1}{2}$ inches long and up to 6 inches wide may be used.

The dotted line at S represents a piece of any opaque

material such as a piece of thin board, that may be placed in front of the lens *L* in order to prevent the beam of light from striking the sensitive film. At *K* is located a key (shown in detail in Fig. 3) to be operated by the subject whose

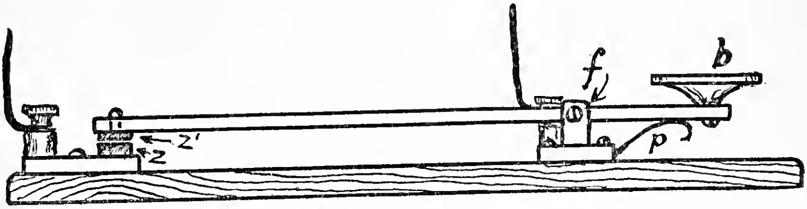


Fig. 3

reaction time is desired. By depressing the key the circuit is instantly opened and the arc extinguished. In order that the circuit may be quickly opened, the key is made as a lever with its longer arm on the opposite side of the fulcrum *f* (see Fig. 3) from the button *b* which the subject presses.

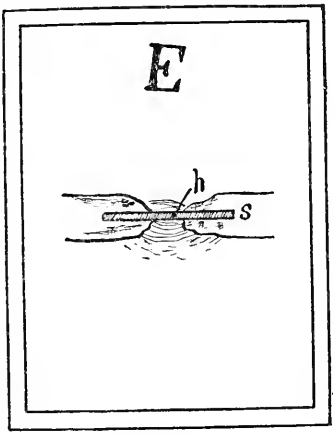


Fig. 4

A slight motion of the button *b* results in a motion at the break end of the arm, four or five times as great as at the button end, thus ensuring a very rapid opening of the circuit. The contacts *Z*, *Z'*, at the break, are small zinc plugs that may be easily removed and replaced as desired.

The screen at *E* (shown in elevation in Fig. 4) has a good

reflecting surface such as white cardboard, and a small slit *s* that allows a portion of the beam from the arc light to pass through a small round hole, *h*, in a movable slider and thence onto the surface of the moving film. The round hole in this slider shields all but a very narrow strip of film from exposure, and allows as many as 70 or 80 records on a film 5 inches wide; the diameter of the hole in the slider being only $\frac{1}{16}$ of an inch in diameter. The hole could be $\frac{1}{32}$ or even $\frac{1}{64}$ in diameter if desired, allowing two or three times as many records on a single film.

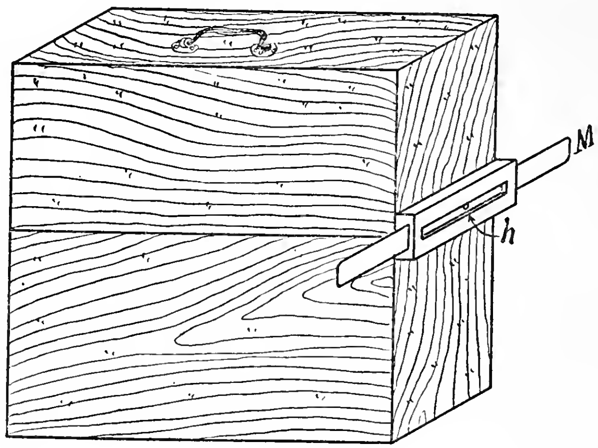


Fig. 5

Fig. 5 shows the wooden case covering the drum, with the movable slider *M* in which is the small hole; shown at *h*. The drum is rotated by means of a $\frac{1}{16}$ horse-power induction motor, properly belted to give the desired speed to the drum. The speed of the drum in making reaction-time records is about one revolution per second.

The experiment for obtaining the reaction-time for sight is performed as follows:

The subject, located at *B*, has the fingers of his hand on the key at *b*, and is told to depress the key as quickly as possible after seeing the image of the arc appear on the screen *E*.

When the opaque screen or shutter *S* is suddenly removed, the image of the arc flashes upon the screen *E* covering a considerable portion of the screen, while a small portion of the light passes through the slit in the screen and the small hole *h* in the slide *M* onto the rapidly moving film. As soon as the subject depresses the key after perceiving the image of the arc, the arc is extinguished and the record on the moving film ceases. The interval of time between the appearance of the image on the screen and the discontinuance of the arc, is the reaction-time of the individual, for sight.

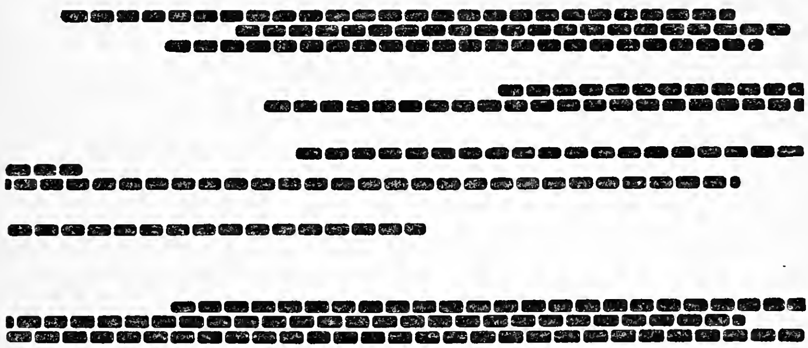


Fig. 6

The shutter *S* is then replaced, the arc again started, by bringing the carbons into momentary contact and then slightly separating them, the slide *M* moved along the width of the hole *h*, and the experiment repeated.

After sufficient records have been made to occupy the width of the film, and the film is removed and developed, blue prints may be made for purposes of reference and computations.

Records are evaluated by counting the number of complete dashes and measuring the fraction of a dash. The indication of the frequency meter is noted for the time each record is made. For accurate results the records should be made in a carefully noted sequence, numbered from a certain side of the film, and the direction of the motion of the film noted.

The method lends itself to class study, since any number

of blue prints may be made and distributed for individual computation and study.

It may be noted that the record is made by a *massless* index, a beam of light that travels about 188,000 miles in a second; that is about 992,640,000 *feet* per second. Since the arc at A is only three or four feet from the film, the time required to reach the film after the shutter S is removed is about 1/248,160,000 of a second; which is small when compared to the actual interval of reaction-time; this being from $\frac{1}{10}$ to $\frac{1}{5}$ of a second.

When the key opens the circuit the current ceases very quickly; whether or not the time of discontinuance of the circuit is of the same order as the velocity of light, it is very small in comparison with the quantity measured.

Fig. 6 is a reproduction from a blue print, of a section of a film obtained from experiment, showing several complete records, using a 60-cycle arc.

VISUAL DISCRIMINATION OF RECTANGULAR AREAS ILLUMINATED BY VARYING DEGREES OF ACHROMATIC LIGHT[1]

BY G. F. ARPS

Ohio State University

The primary purpose of this study is to determine the discriminative efficiency in the perception of areal differences and to analyze the various factors involved in the judgments.

Two series of experiments were made. None of the data for the first series is incorporated in this study, except three graphs, for the reason that as the experiment progressed it became more and more evident that discriminative efficiency, for the illuminations here investigated, is not a simple function of the illuminations. The rejected data (4,212 judgments) comprise one third of the total number of judgments given by all observers.

Means and Methods

The apparatus consists essentially of a light-tight box, a modified form of memory apparatus, tachistoscope, and a contact seconds pendulum which controls the magnet of the tachistoscope.

The entire apparatus is articulated by three independent circuits (.Fig. 1): One alternating circuit, A. C., which provides light for the lamp on the trolley in the light box; one direct current, D. C. P., which enervates the pendulum and makes and breaks a second direct current, D. C. T., which enervates the tachistoscope. By means of a double-knife switch, S, the light in the box is switched off and on

[1] A paper presented before the American Psychological Association at the Chicago Meeting, December, 1915. I am indebted to Mr. C. M. Bock, fellow in the department, for his untiring work as experimenter and for the tabulation of the data. Acknowledgment is made to Kenneth Cottingham for preliminary experimentation and for construction of much of the apparatus.

by a partial depression of the knife; by a complete depression, the tachistoscope is set in function.

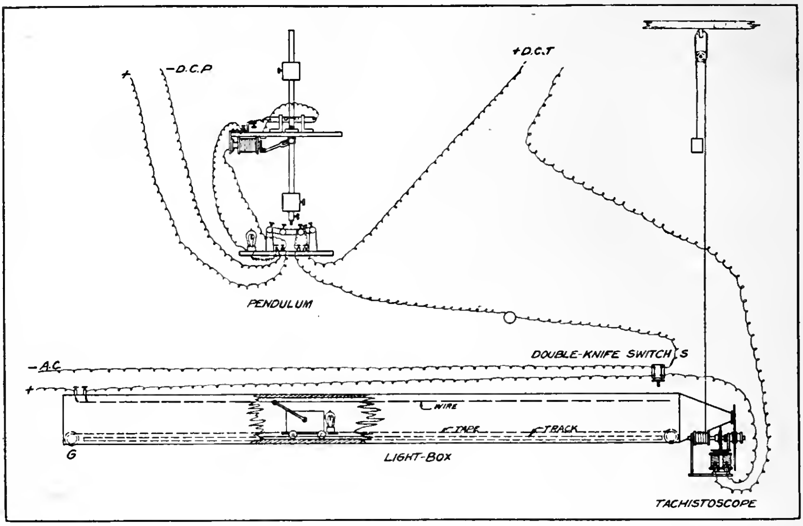


Fig. 1.
Diagram of Apparatus and Electrical Connections.

The box is 14 feet long, 16 inches high, by 12 inches wide, the inside of which is painted black. Within the box is the light carriage, mounted on flanged wheels running the length of the box upon a pair of rails. The light carriage is moved to different distances, depending upon the illumination desired, by a tape which runs, belt-like, through the box and beneath it, passing over a roller, G, at one end of the box.

One end of the box is covered with a tin truncated pyramid, the smaller end of which is pierced by an aperture covered by a piece of ground glass through which the light passes to the tachistoscopic slots. The tachistoscope consists of a circular disk, in whose periphery, placed at equal intervals, are twelve rectangular shaped apertures. Six of these apertures, or slots, are invariable in size and constitute the 'standard' stimulus. The remaining six slots are variable and are provided at one end with a slide by means of which the lengths of the rectangles can be varied at pleasure. These constitute the 'variable' stimuli and are presented

alternately with the standard, the time interval being two seconds. For a cut of the tachistoscope see FIG. 2.

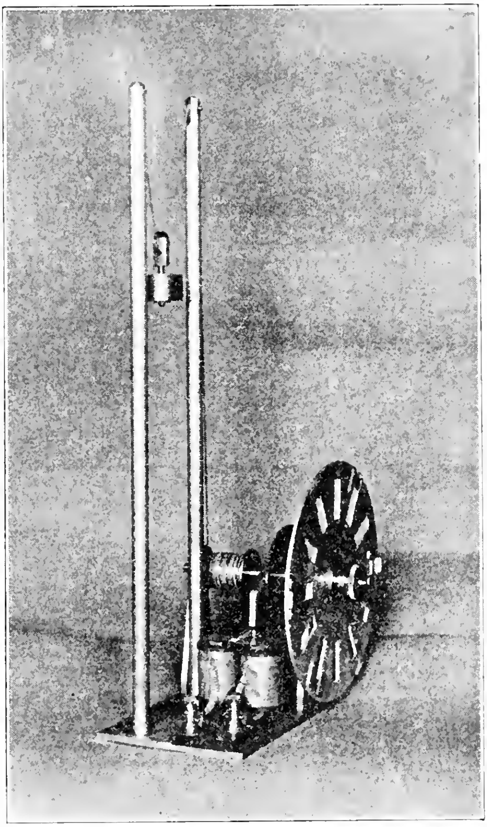


FIG. 2.

The six standard oblongs are exactly 40 mm. in length and 15 mm. in width; the remaining six vary as to length between the limits, lower and upper, of 38.50 and 41.50 millimeters, with steps of one quarter of a millimeter. In this way twelve variations, exclusive of the standard slot, are made possible.

In all, nine illuminations are employed as given in the graphs. An experiment consists in passing before the observer all of the twelve variables, each variable being pre-

ceded by a standard stimulus. An interval of two seconds separates the successive presentation of the standard and variable stimuli. Six judgments are given by each observer, on each variable for each illumination. A more detailed statement of the procedure is as follows: Of the twelve variables, six sizes are chosen by chance, and the oblongs equated on the disk accordingly, the order of the appearance of the oblongs on the disk again being left to chance. The disk is then rotated before the observer, in the manner explained above, and a judgment required on each variable as it appears. The remaining six sizes are now selected and equated as the first six, and similarly used. The division of the twelve variables into two sets, of six each, is made necessary for the reason that the disk was not large enough to contain all the variables and standards at one time.

Following essentially the procedure described for the variables, six judgments are made on the standards as to their variability. The total number of judgments beginning with the shortest or longest variable to, and including, the longest or shortest is, therefore, thirteen. All judgments are expressed as 'longer' (+), 'shorter' (−), or 'equal' (=).

Three observers function in the experiment, *A*, *B*, and *C*. One of the observers is an instructor in the department of psychology and two are graduate students of extended psychological training. The middle point of the illuminated areas is the most favorable fixation point and is maintained throughout.

The above procedure is carried out with the stimulus areas in four positions, viz., the horizontal, vertical and two intermediate positions in which the oblongs incline to the right or left of the vertical position at an angle of 45 degrees. The inclined positions are made possible by tilting the tachistoscope.

Results of the Experiments

(*a*) *Individual Efficiencies*

The important results are set forth, in sixteen graphs, four composite and twelve individual. One additional graph

shows the relative discriminative efficiency, the relative inefficiency and the relative mediocrity of the various illuminations.

The final efficiency value (limen) for each point on the individual graphs is obtained by taking the mean of the negative and positive limens. The negative limen, for example, is computed from a decreasing series of variable stimuli. When two thirds of the judgments on a variable, nearest the standard, are correct, the value of that variable represents the limen of discrimination. The positive values are similarly computed.

The final efficiency value for each individual graph on each illumination is derived from 78 judgments, 6 judgments on each of the 13 variables. The final efficiency value for each illumination on a composite graph for any one of the four positions investigated, therefore, represents a total of 234 judgments. Since there are four positions investigated, each illumination receives 936 judgments.

For the horizontal position, maximal and minimal discriminative efficiency is found at 5.86 and 8.20 c.p. respectively. For the vertical position, maximal efficiency shifts to 1.80 and 4.06 c.p. and minimal efficiency to .50, 1.06 and 2.29 c.p. For the diagonal positions, left and right of the vertical position, maximal efficiency is found between 1.06 and 1.80 c.p. and at 1.06 c.p. respectively, while minimal efficiency for the same positions lies between 2.29 and 8.20 c.p. and at 5.86 c.p. respectively.

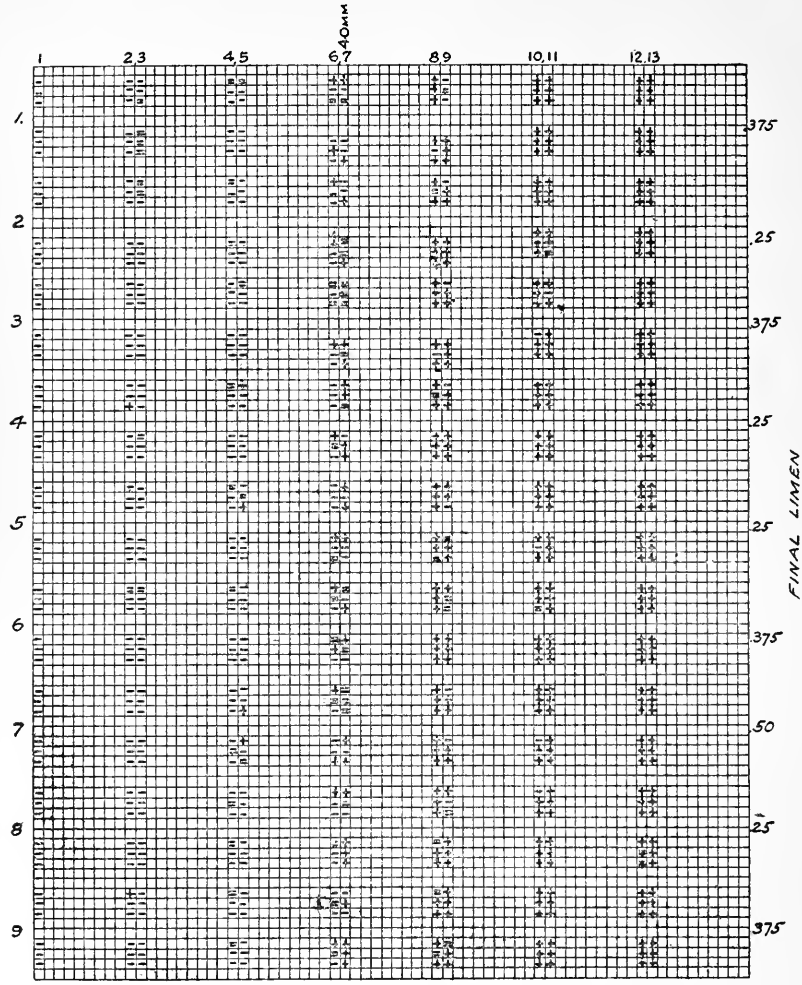
The greatest variations between the maximal and minimal discriminative efficiencies for each of the positions investigated as shown by the composite graphs are as follows: .15 mm. for the horizontal, .080 for the vertical, .20 for the left diagonal, and .080 for the right diagonal position.

Tables I., II. and III., corresponding to graphs II. '*A*,' III. '*B*' and IV. '*C*,' illustrate the distribution of judgments, the relation of the variables, and the negative, positive and final limens.

The character of the data lends itself readily to the computation of the relative efficiency, the relative inefficiency

TABLE I

OBSERVER "A" HORIZONTAL POSITION

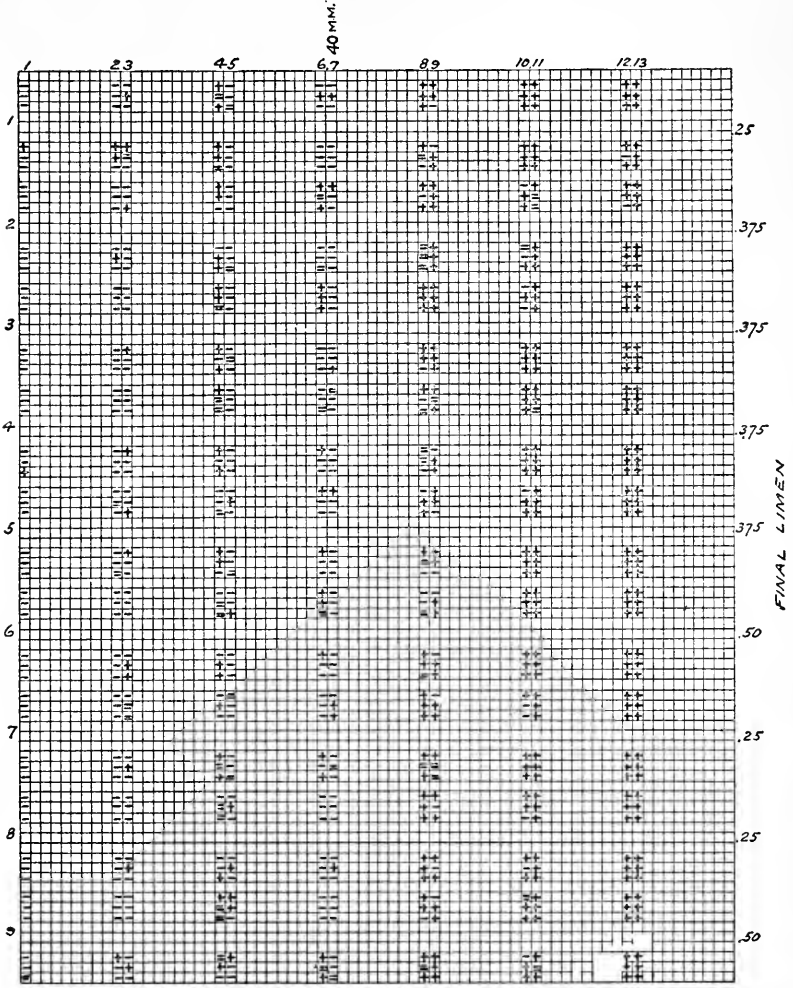


Numerals at the top, 1, 2, 3, etc. indicate the size of the variables, 38.50 mm., 38.75 mm., 39 mm., etc.
" " " " left hand edge, 1, 2, 3, 4, etc. indicate the successive illuminations, .20, .50, 1.06 c.p. etc.

and the relative mediocrity of the various illuminations for all observers. It also lends itself readily to the computation ot the relative efficiency of the observers themselves.

TABLE 2

OBSERVER "B" HORIZONTAL POSITION

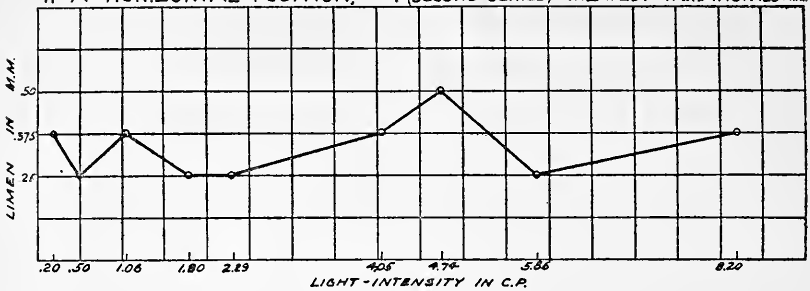


Numerals at the top, 1, 2, 3, etc. indicate the size of the variables, 38.50 mm., 38.75 mm., 39 mm. etc.
" " " " left hand edge, 1, 2, 3, 4, etc. indicate the successive illuminations, 20, .50, 1.06 c.p. etc.

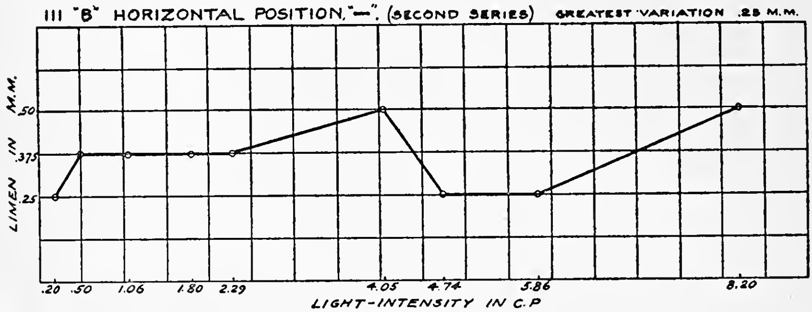
(*b*) *Relative Efficiency of the Illuminations*

The relative efficiency of the various illuminations is derived from a consideration of the low limen of these illuminations. The results show that the illumination of

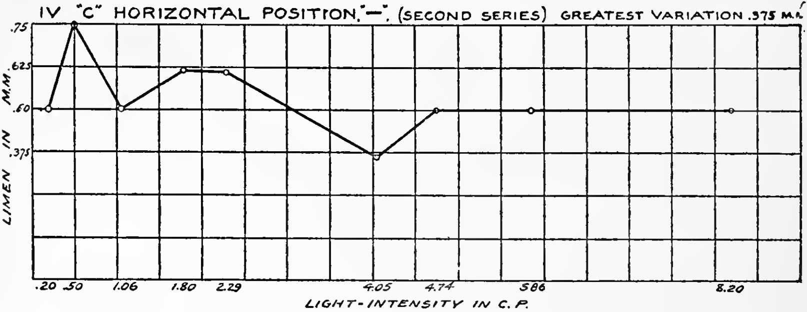
II "A" HORIZONTAL POSITION, "—". (SECOND SERIES) GREATEST VARIATION .25 M.M.



III "B" HORIZONTAL POSITION, "—". (SECOND SERIES) GREATEST VARIATION .25 M.M.



IV "C" HORIZONTAL POSITION, "—". (SECOND SERIES) GREATEST VARIATION .375 M.M.



I COMPOSITE HORIZONTAL POSITION, "—". (SECOND SERIES) GREATEST VARIATION .15 M.M.

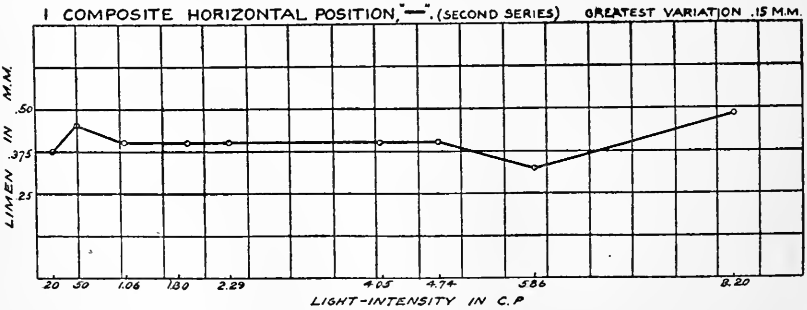
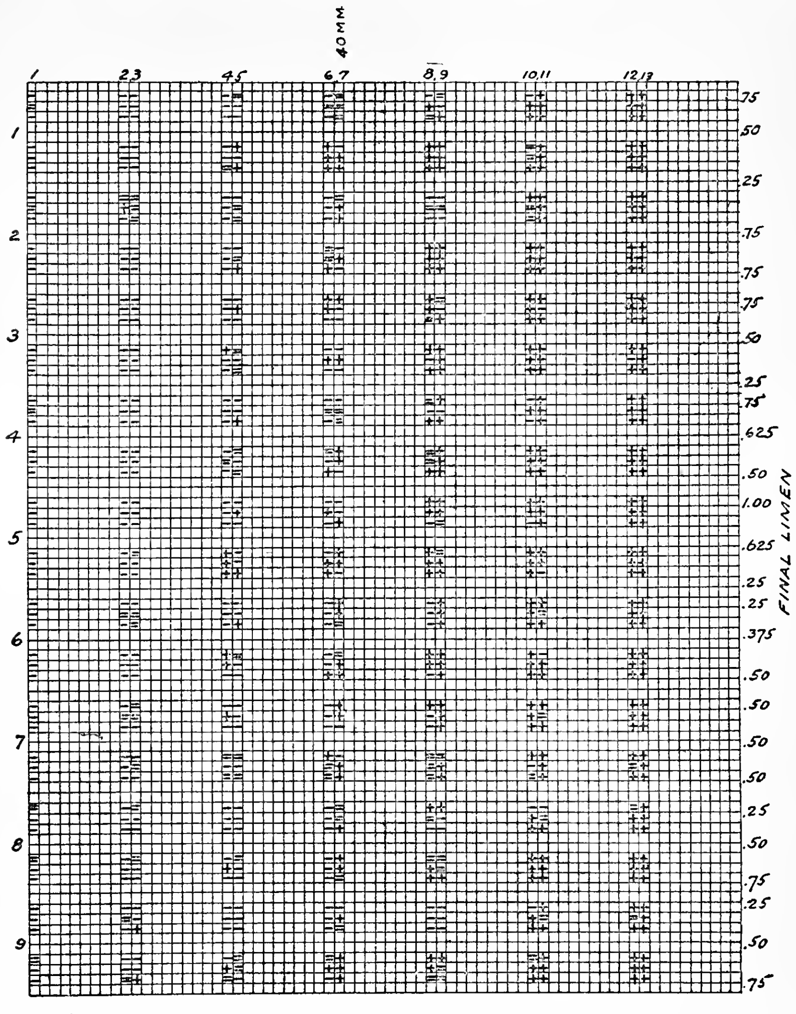


TABLE 3

OBSERVER "C" HORIZONTAL POSITION



Numerals at the top, 1, 2,3, etc. indicate the size of the variables, 38.50 m.m., 38.75 m.m. 39 m.m. etc.
" " " left hand edge, 1, 2,3,4, etc. indicate the successive illuminations, 20, 50, 1.06 c.p. etc.

1.06 c.p. has the highest relative efficiency value of all the illuminations, being 28 on a scale ranging from 1 to 30, while the illumination 2.29 c.p. has the lowest value, being 4 on the same scale.

Deriving now the relative *in*efficiency of the various illuminations through a consideration of the high limen we find that illuminations .50 and 4.74 c.p. are equally and

Table IV

Résumé and Comparative Tables—Observers *A*, *B*, *C*

'Figures' are in 'thousands' of one millimeter, and indicate the average limen, as defined in this study.

From the columns *A*, *B*, *C*, are derived the credit values given under Table V.

	C. P.	*A*	*B*	*C*	Av.
'—'	.20	375	250	500	375
	.50	250	375	750	465
	1.06	375	375	500	416
	1.80	250	375	625	416
	2.29	250	375	625	416
	4.06	375	500	375	416
	4.74	500	250	500	416
	5.86	250	250	500	333
	8.20	375	500	500	486
					Average, 416
'/'	.20	500	375	500	458
	.50	625	500	250	458
	1.06	500	250	250	333
	1.80	625	375	250	420
	2.29	625	500	375	500
	4.06	500	500	375	458
	4.74	625	375	500	500
	5.86	625	500	500	540
	8.20	500	500	375	458
					Average, 458
'\'	.20	625	500	375	500
	.50	625	500	375	500
	1.06	500	375	500	458
	1.80	500	375	500	458
	2.29	625	500	500	542
	4.06	625	500	500	542
	4.74	625	500	500	542
	5.86	625	500	500	542
	8.20	625	500	500	542
					Average, 514
'|'	.20	375	500	250	375
	.50	500	500	250	413
	1.06	500	500	250	413
	1.80	375	375	250	333
	2.29	500	500	250	413
	4.06	375	375	250	333
	4.74	500	375	250	375
	5.86	500	375	250	375
	8.20	375	375	250	333
					Average, 373

most inefficient relatively, while illumination 1.80 c.p. is, relatively, least inefficient of all the illuminations.

Finally, if we consider the relative mediocrity of all the illuminations by a consideration of the average limen we find that illuminations .20 and .50 c.p. have the highest and lowest mediocre value respectively (Tables IV. and V.).

All the above relationships are set forth in three curves (Graph XVII.), derived from the data summarized in Table VI.

Table V

Relative Efficiency Coefficients of Various Illuminations for All Observers

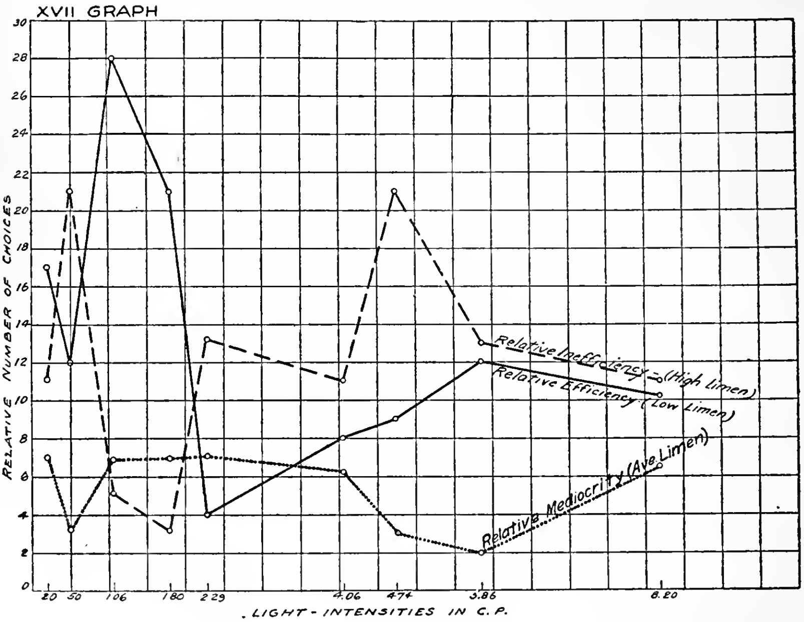
Illumination	—A	—B	—C	A—/	B—/	C—/	A—\	B—\	C—\	A—|	B—|	C—|	Total
.20		333	200	250					500	250		111	1.644
.50	250					333			500			111	1.194
1.06			200	250	1.000	333	500	500				111	2.894
1.80	250					333	500	500		250	200	111	2.144
Low limen, 2.29	250											111	.361
4.06				250						250	200	111	.811
4.74		333	200								200	111	.844
5.86	250	333	200								200	111	1.299
8.20			200	250						250	200	111	1.011
.20	250		140		333								.723
.50		250											.250
1.06	250	250	140										.640
1.80		250	140		333								.723
Av. limen, 2.29		250	140			333							.723
4.06	250					333							.583
4.74			140		333								.473
5.86			140										.140
8.20	250		140			333							.723
.20						333	140	140			250		1.063
.50			1.000	200	200		140	140		200	250		2.130
1.06									140	200	250		.590
1.80				200					140				.340
High limen, 2.29				200	200		140	140	140	200	250		1.270
4.06		500			200		140	140	140				1.120
4.74	1.000			200		333	140	140	140	200			2.153
5.86				200	200	333	140	140	140	200			1.353
8.20		500			200		140	140	140				1.120

Method.—For any position and any observer select that illumination which has the lowest limen; to this illumination, give one unit credit. Do this for all positions and all observers. Summate. These sums are the relative efficiency coefficient. In case two or more illuminations share the low limen, give each one half, etc., credit. For the derivation of the credit values given above see Table IV.

Table VI

Summary of Table V

Illumination	Gets Low Limen	Average Limen	High Limen
.20	1,644 times	723 times	1,063 times
.50	1,194 "	250 "	2,130 "
1.06	2,894 "	640 "	590 "
1.80	2,144 "	723 "	340 "
2.29	361 "	723 "	1,270 "
4.06	811 "	583 "	1,120 "
4.74	844 "	473 "	2,153 "
5.86	1,229 "	723 "	1,353 "
8.20	1,011 "	723 "	1,120 "



(c) *Relative Efficiency of the Observers*

To determine the relative efficiency of the observers, each observer was ranked in order of his limen for all illuminations of a given position. The observer having the lowest limen for any given illumination was graded '1'; the one having the highest limen was, therefore, graded '3.' A perfect score for any position would be represented by '9,' since there are 9 illuminations. For example, observer *A* for the horizontal position ranked '1' in 7 of the 9 illuminations

and '2' in 2 of them. If we divide the number of illuminations (9) by the sum of *A*'s ranks, which is 11, we get *A*'s relative efficiency for the horizontal position, which is 0.81. Now if we similarly rank observer *A* for the remaining positions and divide by the number of positions, we get *A*'s relative efficiency for all positions and all illuminations. In a similar manner compute the relative efficiencies of *B* and *C*. The results show that the relative efficiencies of observers *A*, *B*, and *C* are .51, .63, .76 respectively where unity represents perfect efficiency (Table VII.).

TABLE VII

COMPARATIVE TABLE OF OBSERVERS' EFFICIENCY

	Σ	$\frac{—}{\div 9}$	/	\	|	Σ ÷ 4
Horizontal, '—,' Position:						
A 2, 1, 1, 1, 1, 1, 2, 1, 1	11	0.81	0.39	.410	.410	*.510*
B 1, 2, 1, 2, 2, 2, 1, 1, 2	14	0.63	0.63	.810	.460	*.630*
C 3, 3, 2, 3, 3, 1, 2, 2, 2	21	0.43	0.81	.810	1.000	*.760*
Diagonal, '/,' Position:						
A 2, 3, 2, 3, 3, 2, 3, 2, 3	23	0.39				
B 1, 2, 1, 2, 2, 2, 1, 1, 2	14	0.63				
C 2, 1, 1, 1, 1, 1, 2, 1, 1	11	0.81				
Diagonal, '\,' Position:						
A 3, 3, 3, 3, 2, 2, 2, 2, 2	22	0.41				
B 2, 2, 1, 1, 1, 1, 1, 1, 1	11	0.81				
C 1, 1, 2, 2, 1, 1, 1, 1, 1	11	0.81				
Vertical, '|,' Position:						
A 2, 2, 2, 2, 2, 2, 3, 3, 2	20	0.41				
B 3, 2, 2, 2, 2, 2, 2, 2, 2	19	0.460				
C 1, 1, 1, 1, 1, 1, 1, 1, 1	9	1.000				

Method.—Rank the observers in order of their limens, for all illuminations of a given position. Summate. Divide each such sum by '9,' since that would have been a perfect score, relatively, if any observer had stood first in all the illuminations. Do this for other positions, summate, and divide by '4' to get average for all positions.

(*d*) *Relative Efficiency of the Positions*

Of the four positions, the horizontal and vertical positions are equal in relative efficiency and both higher than the diagonal positions. The right slanting diagonal stands next in efficiency while the left slanting diagonal is least efficient. The numerical equivalents of the various positions are: .59,

.47, .41, .59 respectively for the horizontal, right diagonal, left diagonal and vertical. On the basis of normal eye usage we should expect the order of relative efficiency to follow

Table VIII

Résumé and Comparative Table Showing Relative Efficiency of the Four Positions

Obs.	Limen in Thousands of a Mm.					Rank	Coefficient of Efficiency
	'—'	'/'	'\'	'|'			
A	375	500	625	375	.20	1, 2, 3, 1	
B	250	375	500	500	.20	1, 2, 3, 3	
C	500	500	375	250	.20	3, 3, 2, 1	
A	250	625	625	500	.50	1, 3, 3, 2	
B	375	500	500	500	.50	1, 2, 2, 2	
C	750	250	375	250	.50	3, 1, 2, 1	
A	375	500	500	500	1.06	1, 2, 2, 2	
B	375	250	375	500	1.06	2, 1, 2, 3	
C	500	250	500	250	1.06	2, 1, 2, 1	
A	250	625	500	375	1.80	1, 4, 3, 2	
B	375	375	375	375	1.80	1, 1, 1, 1	
C	625	250	500	250	1.80	3, 1, 2, 1	
A	250	625	625	500	2.29	1, 3, 3, 2	
B	375	500	500	500	2.29	1, 2, 2, 2	
C	625	375	500	250	2.29	4, 2, 3, 1	
A	375	500	625	375	4.06	1, 2, 3, 1	
B	500	500	500	375	4.06	2, 2, 2, 1	
C	375	375	500	250	4.06	3, 3, 2, 1	
A	500	625	625	500	4.74	1, 2, 2, 1	
B	250	375	500	375	4.74	1, 2, 3, 2	
C	500	500	500	250	4.74	2, 2, 2, 1	
A	250	625	625	500	5.86	1, 3, 3, 2	
B	250	500	500	375	5.86	1, 3, 3, 2	
C	500	500	500	250	5.86	2, 2, 2, 1	
A	375	500	625	375	8.20	1, 2, 3, 1	
B	500	500	500	375	8.20	2, 2, 2, 1	
C	500	375	500	250	8.20	3, 2, 3, 1	
Total '—'						46	.59 (27/46)
" '/'						57	.47 (27/57)
" '\'						65	.41 (27/65)
" '|'						40	.59 (27/40)

Method.—Give to that position, in the column marked 'Rank' a number, to indicate whether that position had the lowest, second lowest, etc., limen. Add the four vertical columns of figures. Divide 27 by each such sum since any position, had it had the lowest limen all the time, would have polled $27 \times 1 = 27$.

what we have found, with the exception that the vertical position should be next in efficiency to the horizontal but not equal to it (see Table VIII.).

Observations and Conclusions

An analysis of certain of the factors involved in the judgments of spacial differences of the areas here investigated yields some interesting results.

1. So far as could be determined, eye-movement appears to play no direct role in a judgment. The chief function of movement in the 'free' eye is to secure and maintain a position maximally favorable to adequate retinal stimulation. Conscious increment of eye-movement acts certainly as a distraction. Observer *A* says: "When I run my eyes over the oblong, horizontally and vertically, in order better to grasp the area, I lose all sense of security. At such times I frequently request the experimenter to repeat the series. On the other hand, inhibition of eye-movement, through fixation of the middle point of a rectangle, increases the sense of security and accuracy."

The chief criterion for a judgment in the case of observer *A* appears to be the total feeling of largeness or smallness of the compared areas, and this is true whether the rectangle as a whole (by fixating a middle point) or one of the lateral lines (by fixating the middle point of the line) is perceived. In confirmation of this fact several observations were made by fixating the middle point of one of the lateral lines. This 'total feeling' is correlated with the degree of simultaneous stimulation of retinal points. The kinæsthetic element due to eye-movement is not considered the primary criterion in the judgments of extents here investigated.

Observers *B* and *C* experience a peculiar feeling of 'thickening' or 'thinning' of a comparative oblong according to whether the judgment is 'longer' or 'shorter.' This illusion is an invariable criterion for a judgment and appears constant. Observer *C* says: "I get the feeling of size best when I maintain an unmoved attitude of completeness, when I take in the rectangle as a whole for then the illusory effect of 'thin-

ning' and 'thickening' is pronounced. I do better when I get a total impression of eye-stimulation."

2. The exact function of the standard stimulus in the judgment of a variable is difficult of definition. Observer *C* says: "When the variable is in the field I appear to judge it independently on its own merits, without the slightest back reference to the standard. I appear to have no after-image of the standard so as to make superposition possible." What lingering effects the standard stimulus may bring to the presence of the variable seems indissolubly merged with the latter. The judgment of a variable as 'longer' or 'shorter' rests apparently on the absolute value of the variable, and is immediately perceived as an expansion or shrinkage of the variable. Martin and Muller[1] and other investigators[2] in other fields have made similar observations.

3. During the entire period of experimentation all observers protested the real size of the standard area. Frequently at the close of a given series observer *A* gave it as his opinion that the normal sequence (standard followed by a variable) was not adhered to and that all of the judgments were consequently not made on the variable. Closer inquiry shows conclusively that the standard at times varied subjectively, that the same retinal impressions permitted of ambiguous interpretation. In this illusion the standard area altered its apparent size without a corresponding change in distance. This was, no doubt, true of the variables, but in such cases it was difficult of detection for the reason that a variable was expected to alter its size. Since the illusion is neither constant nor necessarily restricted to the standard area, it is apparent that such fluctuations have an important bearing on the accuracy of judgment. The partial control of this illusion and the reduction of the effects of practice may account for the almost horizontal appearance of the graphs. On the other hand, the inadvertent appearance of the illu-

[1] *Beitrage zur Analyse der Unterschiedsempfindlichkeit*, pp. 44–45.

[2] Arps, 'Introspective Analysis of Certain Tactual Phenomena,' *Psychological Review*, Vol. 19, pp. 345–346. Stumpf and Meyer, *Zeitschrift für Psychologie der Sinnesorgane*, Vol. 18, pp. 390–392.

sion may account for the different positions of the limens on the individual graphs.

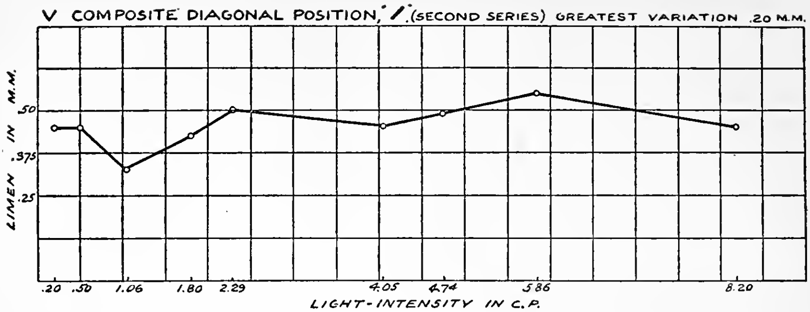
With reference to the illusion it should be added that when the fixation point is beyond the stimulus area we should expect double images of the area itself to appear. None of the observers, however, report the presence of such images.

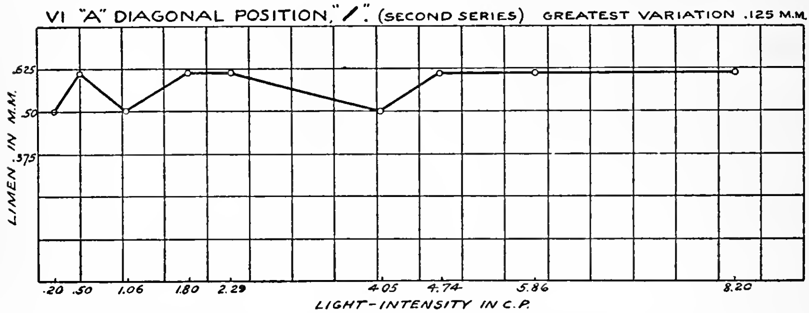
The explanation of the illusion is simple. Usually a change in size parallels a corresponding change in distance. Likewise a change in distance parallels a change in the optic axes. It is probable that we have here inadequate convergence and that the optic axes do not intersect at the stimulus area but beyond, so that the axes are more nearly parallel than is the case when the stimulus area is in focus. Thus the feeling of convergence is less intense than experience teaches is necessary to perceive the area as such a size and at such a distance. "If degree of convergence is a criterion for distance and if distance is a measure for the apparent size of an object then we have all the conditions necessary for an explanation of the illusion."[1]

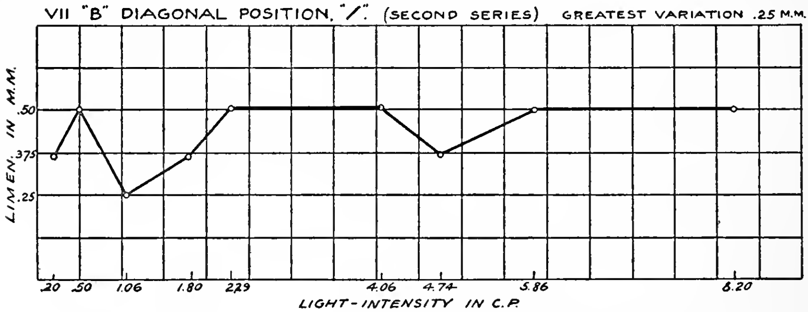
Connected with the apparent change in size of certain of the standard stimuli is the illusion of distance. Both observers *A* and *C* report that whenever the illusion of size is marked, the oblong appears to recede. Observer *A* says: "Whenever the standard varies decidedly, it seems more remote; at such times the succeeding variable fairly snaps into a nearer position."

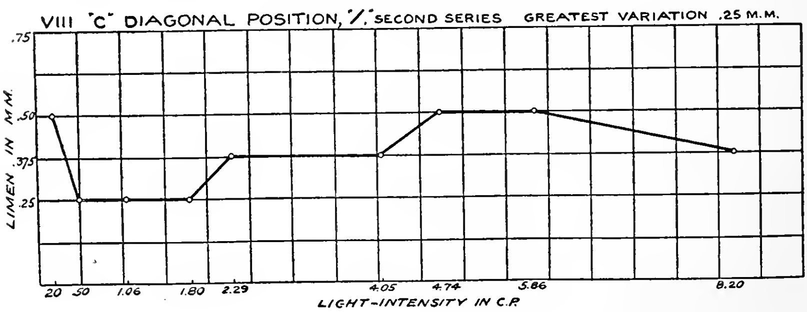
Here again the explanation is simple. Size and distance through innumerable repetition in experience, are linked into a permanent association, so that the one habitually becomes a sign for the other. A change in distance or size, therefore, is followed by the associated change in size or distance. For this reason an apparent decrease in the size of the standard oblong is supplemented by the idea of the oblong's greater distance. The two-dimensional space perception of the oblong is, consequently, enlarged into a tri-dimensional perception in the case of the standard oblongs with which the illusion of size is connected.

[1] Arps, 'Two Interesting Cases of Illusion of Perception,' *Journal of Abnormal Psychology*, Vol. X., pp. 211–212.

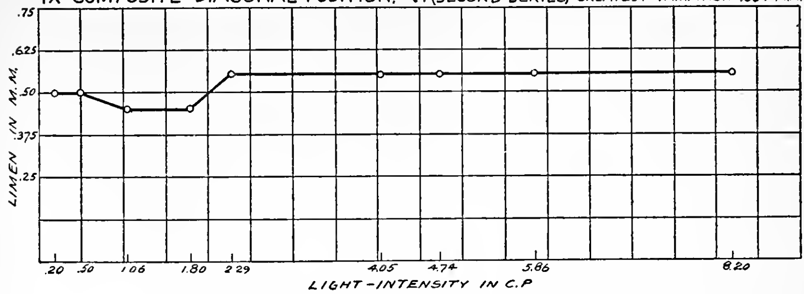

V COMPOSITE DIAGONAL POSITION, "/." (SECOND SERIES) GREATEST VARIATION .20 M.M.
LIMEN IN M.M.
.50
.375
.25
.20 .50 1.06 1.80 2.29 4.05 4.74 5.86 8.20
LIGHT-INTENSITY IN C.P.


VI "A" DIAGONAL POSITION, "/." (SECOND SERIES) GREATEST VARIATION .125 M.M.
LIMEN IN M.M.
.625
.50
.375
.20 .50 1.06 1.80 2.29 4.05 4.74 5.86 8.20
LIGHT-INTENSITY IN C.P.

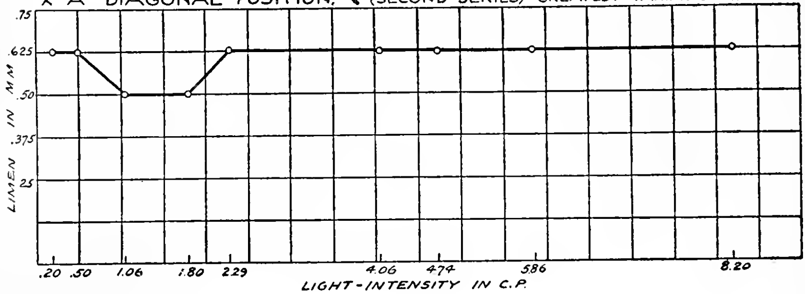

VII "B" DIAGONAL POSITION, "/." (SECOND SERIES) GREATEST VARIATION .25 M.M.
LIMEN IN M.M.
.50
.375
.25
.20 .50 1.06 1.80 2.29 4.06 4.74 5.86 8.20
LIGHT-INTENSITY IN C.P.


VIII "C" DIAGONAL POSITION, "/." SECOND SERIES GREATEST VARIATION .25 M.M.
LIMEN IN M.M.
.75
.50
.375
.25
20 .50 1.06 1.80 2.29 4.05 4.74 5.86 8.20
LIGHT-INTENSITY IN C.P.

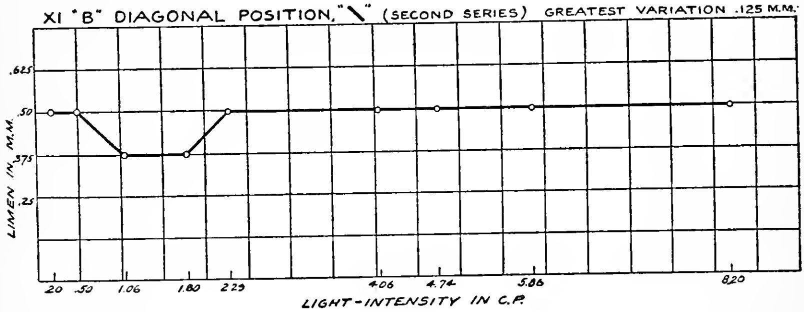
IX COMPOSITE DIAGONAL POSITION, "\". (SECOND SERIES) GREATEST VARIATION .084 M.M.



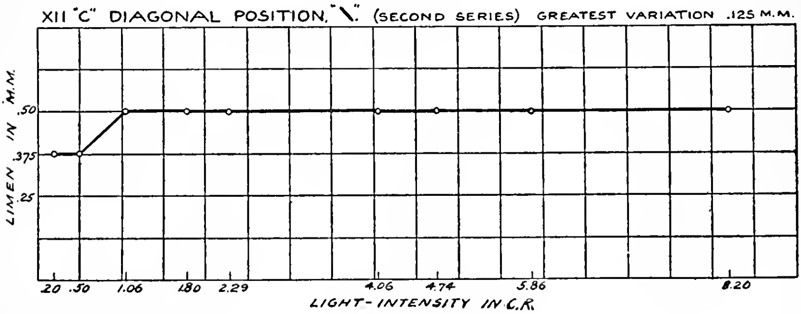
X "A" DIAGONAL POSITION, "\" (SECOND SERIES) GREATEST VARIATION .125 M.M.

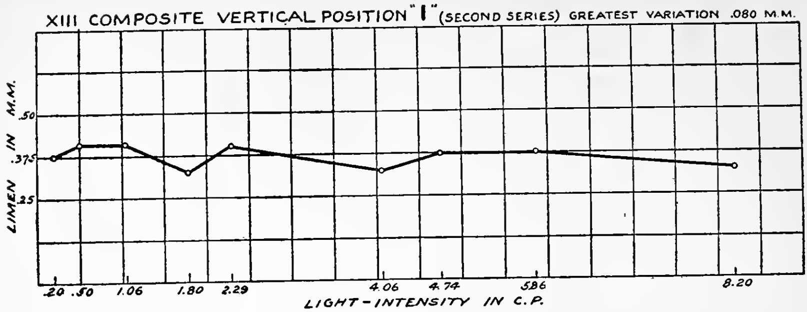


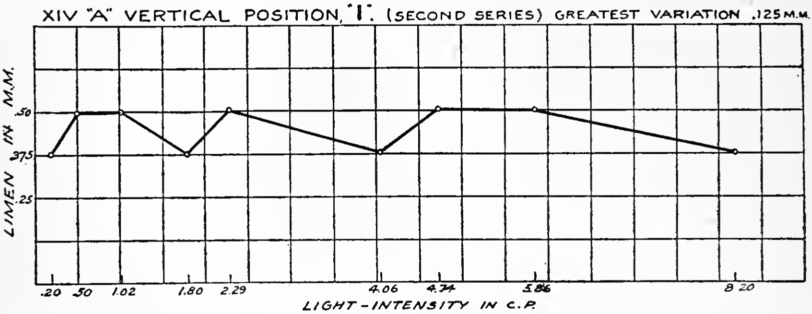
XI "B" DIAGONAL POSITION, "\" (SECOND SERIES) GREATEST VARIATION .125 M.M.

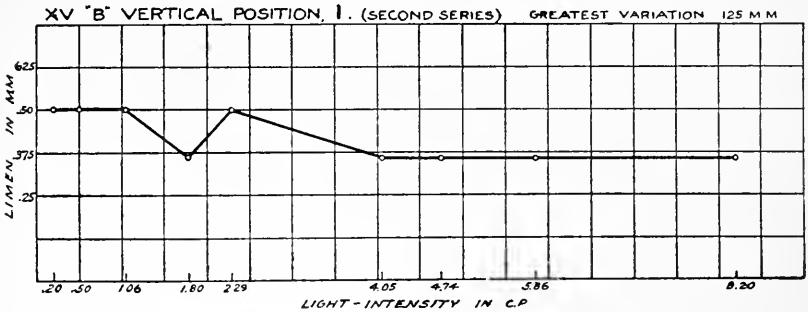


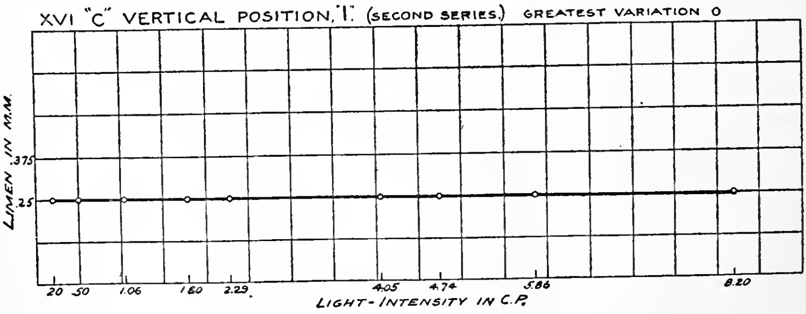
XII "C" DIAGONAL POSITION, "\" (SECOND SERIES) GREATEST VARIATION .125 M.M.




XIII COMPOSITE VERTICAL POSITION "I" (SECOND SERIES) GREATEST VARIATION .080 M.M.
LIMEN IN M.M.
.50
.375
.25
.20 .50 1.06 1.80 2.29 4.06 4.74 5.86 8.20
LIGHT-INTENSITY IN C.P.

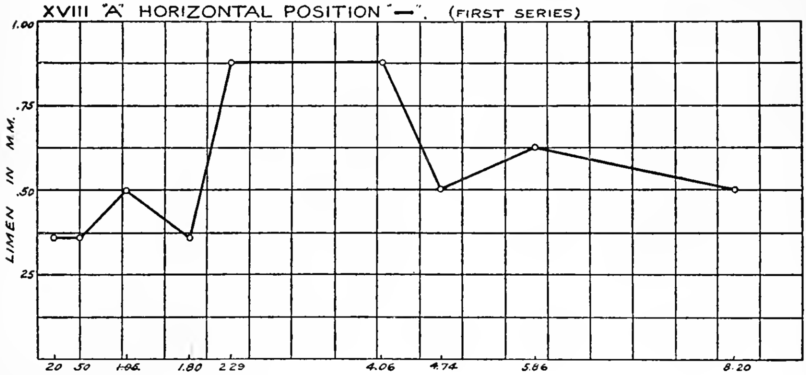

XIV "A" VERTICAL POSITION, "I". (SECOND SERIES) GREATEST VARIATION .125 M.M.
LIMEN IN M.M.
.50
.375
.25
.20 .50 1.02 1.80 2.29 4.06 4.74 5.86 8.20
LIGHT-INTENSITY IN C.P.


XV "B" VERTICAL POSITION, I. (SECOND SERIES) GREATEST VARIATION 125 M M
LIMEN IN MM
.625
.50
.375
.25
.20 .50 1.06 1.80 2.29 4.05 4.74 5.86 8.20
LIGHT-INTENSITY IN C.P


XVI "C" VERTICAL POSITION, "I". (SECOND SERIES.) GREATEST VARIATION 0
LIMEN IN M.M.
.375
.25
.20 .50 1.06 1.80 2.29 4.05 4.74 5.86 8.20
LIGHT-INTENSITY IN C.P.

4. The graphs showing the relative efficiency, the relative inefficiency, the relative mediocrity of the illuminations and the individual and composite graphs seem to show that the limen is not a simple function of the illumination, for then we would expect a curve with a final efficiency maximum, *i. e.*, we would expect one illumination to have the greatest efficiency for all observers.

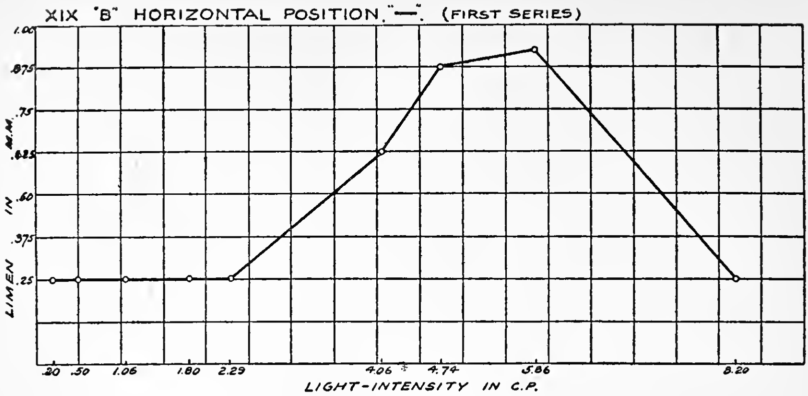
XVIII "A" HORIZONTAL POSITION "—". (FIRST SERIES)

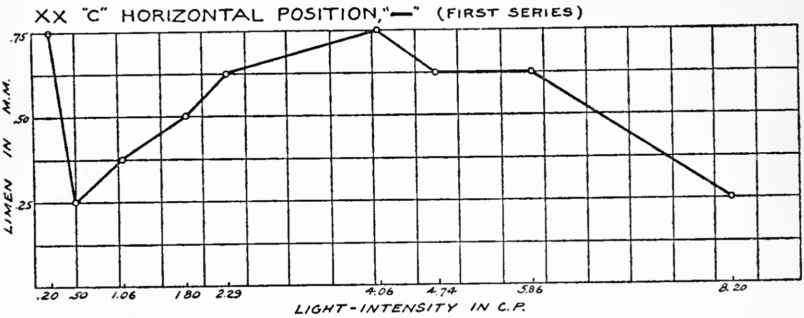


In case the illumination at the lowest limit were reduced in intensity to the point of bare perceptibility, and, at the highest limit increased to a disagreeable, or, even painful degree, we would expect very high limens at these two extremes of illumination. Convincing evidence on this point must be left, of course, for future experimentation.

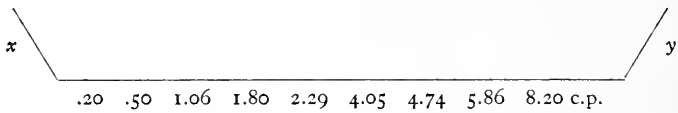
When we compare the graphs XVIII. '*A*,' XIX. '*B*' and XX. '*C*' of the first series of experiments with those of the second series, the general ironing out process due to practice is obvious. It is altogether likely that discriminative efficiency which determines the character of the graphs of the second series of experiments *is a simple function of habit.* It is not unlikely that the individual graphs would assume a straight line if it were possible to eliminate the illusion mentioned above, and certain unavoidable errors as may be due, for example, to bodily conditions. The efficiency of observer *C* for the vertical position (Graph XVI. '*C*') seems to substantiate this conclusion.

Taking observer *C*'s efficiency for the vertical position





and assuming high limens for the hypothetical illuminations at the extremes of the series of illuminations, we get the following diagrammatic graph:



It is altogether likely that the relation of the optimal stimuli (.20 c.p. . . . 8.20 c.p.) to the extreme stimuli (x and y) of this series would be typical of the relation existing between the multitudinous visual stimuli of normal life in which the extremes of the artificial series approximately correspond to growing darkness and brilliant sunlight.

Note.—It is desirable to continue the experiment and reduce the differential steps of the variable below .25 mm. The withdrawal of two of the observers, however, makes this impossible, and obviously new recruits can not begin where practiced observers end. Moreover, it is doubtful whether the main conclusions would be affected by the additional data.

AUDITORY ILLUSIONS OF MOVEMENT—A PRELIMINARY STUDY

BY HAROLD E. BURTT

Harvard University

Introduction

The following study was suggested by the experiments of Wertheimer[1] and Korte[2] on visual illusions of movement and of Benussi[3] on similar tactual illusions.

Wertheimer exposed in quick succession two parallel lines a few centimeters apart. Under certain conditions of time and distance it was possible to produce the impression of a line moving in the same direction as the actual temporal succession of the stimuli (the ordinary illusion of the stroboscope and moving-picture). Wertheimer attributed the phenomena to a 'physiologische Kurzschluss' between two regions of the cortex corresponding to the two points in visual space. Korte repeated Wertheimer's work and, carrying the problem still farther, subscribed to the same theory. Benussi, working with two successive pressure stimuli, found somewhat similar results and also favored Wertheimer's theory.

It would seem interesting, inasmuch as the visual and tactual are fields in which there is some plausibility for a correspondence between points of objective space and points in the cortex, to apply the same methods in another sense-department such as audition, where such a correspondence seems highly improbable. Accordingly the present study

[1] Wertheimer, M., 'Experimentelle Studien über das Sehen von Bewegung,' *Zeitschrift für Psychologie*, 1912, 61, pp. 161–265.

[2] Korte, A., 'Kinematoskopische Untersuchungen,' *Zeitschrift für Psychologie*, 1915, 72, pp. 193–296.

[3] Benussi, V., 'Kinematohaptische Erscheinungen,' *Archiv für die gesamte Psychologie*, 1913, 29, pp. 385–388. Benussi, V., 'Kinematohaptische Scheinbewegungen und Auffassungsumformung,' *Bericht VI Kongres für experimentelle Psychologie*, Göttingen, 1914, pp. 31–35.

attempted to repeat the work of Wertheimer and Korte in certain of its aspects in audition.

These investigators found, for instance, that an interval of about 60 sigma between stimuli gave, on the whole, the optimal impression of movement. With longer exposure of the lines, however, the optimal interval was relatively somewhat shorter. Moreover, by increasing the intensity of the second stimulus or by voluntarily directing the attention to it, apparent movement was frequently produced in the reverse direction. Further studies were made with different combinations of lines and with variations of distance and intensity, but only three problems were taken up in the present case:

1. The possibility of auditory illusions of movement.
2. The relation between exposure and interval.
3. The effect of difference in the intensity of the two stimuli.

Technical difficulties made it inadvisable to proceed farther with the present form of apparatus.

Apparatus and Method

The stimuli consisted of telephone receivers in series with a 250 vd. tuning fork on a 5-volt circuit. Normally the current passed through the fork directly. When, however, one side of the line was broken by relays actuated by a time-controlling mechanism the current was shunted through one of the receivers which gave a faint buzzing sound of the same pitch as the fork. One receiver was mounted on an arm moving synchronously with the time-mechanism. Two others were stationary. It was thus possible to produce an actually moving sound, a moving one with a break in the middle, or two separate successive sounds in different positions. For the first two cases it was merely necessary to close switches so that a single relay was operated once or twice, the current in each case passing through the moving receiver. For the two stationary receivers connections were made so that when one relay broke the current passed from the fork through the one receiver and when a second relay broke, the current

passed through the other receiver. Rheostats in series with all the receivers regulated the intensity of the stimuli.

A wooden disc, mounted upright, was rotated by an alternating-current motor reduced by belt gear and speed reducers. Various speeds were obtained by different combinations of gearing. Four brass contacts were clamped over the edge of this disc and held in place by set-screws on the rear side. Nearer the center were four concentric brass rings a few millimeters apart. These were connected through the disc by means of flexible wire to the rear of the four brass contacts. Brushes of spring brass pressed upon these rings, thus affording connections with the four contacts while the disc was rotating. A common pole of spring brass on one side of the main line brushed across the four contacts in succession. These latter passed to various connections operating the magnets of the relays. A telegraph key in series with the main brush enabled the whole line to be thrown on or off at any instant.

In some cases only two contacts were used, the width of the contacts giving the length of time the relay was broken, and hence the length of the sound, and the distance between the contacts giving the interval between the two sounds. For a wider range of adjustments a relay had the spring removed and a second set of magnets on the opposite side of the armature adjusted to just touch the armature when it was in its normal position. The residual magnetism of the cores held the armature there even though the relay was tipped slightly so that the armature tended to fall. The first contact on the disc (when the main brush passed across it) actuated the original magnets, pulling the armature down. It stayed there until the second contact actuated the other magnets pulling the armature back and holding it there by residual magnetism. The third and fourth contacts produced the same effect on a second relay arranged in the same manner. Thus by varying the distance between contacts and the speed of the disc, any desired exposure or interval could be obtained. The exposures and intervals were roughly calibrated by sending the alternating current with lamp resistance through

the relays and a signal magnet writing on a kymograph. This could be read accurately to 1/120th of a second. In some of the work the Hipp chronoscope was also used to measure the time.

The experiments were performed in the sound-proof room of the laboratory. The subject and the stimuli were in an inner room behind double doors. The stationary receivers were hung on the edge of a sound cage one meter in diameter. The subject sat either in the center of the cage or at a distance of two meters from the receivers. In some of the work a special rack with a black cambric screen between the receivers and the subject was used. The moving receiver was fastened to a wooden arm, the lower end of which was pivoted to a rod attached to the table edge. A horizontal rod pivoted to this arm passed through a small hole in the wall and was connected to a shaft operated by a crank attached to the rotating disc of the time-controlling mechanism. Thus the receiver oscillated synchronously with the exposures of the sounds through a distance of 70 cm. The receivers were all on the level of the ears of the subject who sat facing them on an adjustable chair. Preliminary experiments with the subject facing in various directions appeared to yield no significant facts except the more frequent confusion of direction of motion when the sounds were directly at one side, *i. e.*, at the end of the first or third quadrants, considering the eyes as facing zero. This would correspond to the usual confusion in the direction of sounds localized symmetrically with reference to the plane passing vertically through the two ears. The motor and tuning fork were in padded boxes so that the sound was minimized and what passed through into the inner room afforded little distraction but served rather as a warning signal.

At the beginning of the hour's work the receivers were equated in intensity according to the judgment of the subject. He was then simply told that when the light was extinguished he was to give his attention in the direction of the receivers and after the doors were again opened to describe what he had heard. The doors were then closed,

the fork started, the light extinguished and the time mechanism set in operation. Some time between the passing of the fourth and first contacts across the common brush the telegraph key was depressed and held down until the fourth contact had again passed, thus giving the stimuli in the desired arrangement. This arrangement was varied from day to day, sometimes continuously moving sounds intermixed with broken movement, sometimes moving sounds mixed with stationary ones.

Five subjects participated in the experiment, all graduate students with psychological training (one a woman). Only one was at all familiar with the problem and none were informed of the trend of results through the work, although all realized the difficulty of judging and were curious as to their efficiency. The experiments were performed in the Harvard Psychological Laboratory in the first half of the college year 1915–16.[1]

The present study is only of a preliminary nature. The stimuli were far from satisfactory. Although the intensity of the receivers could be adjusted by the rheostats there were invariably differences in timbre. These could be regulated to some extent by screwing the cap which confined the edges of the diaphragm. But undoubtedly the subjects almost always had timbre differences to assist them in distinguishing the discreteness of the two sounds, and hence the illusion took place less frequently and consistently. It would seem desirable to do more exhaustive work with tandem-driven tuning forks in boxes with small apertures, controlling the intensity by rotating the forks.

Results

The principal results for experiments with two sounds of equal intensity are summarized in Table I. The main interest lies, of course, in the trials in which two successive stationary sounds in different places yielded an impression of movement. The actually moving sounds and the broken moving sounds were introduced mainly as a check and to

[1] The writer expresses his obligations to the late Hugo Münsterberg.

avoid the subject's anticipating the character of the stimulus. These are not included in the table. The first few columns give the conditions investigated which yielded the optimal impression of movement for each subject; the next group those conditions in which the two sounds were always recognized as discrete and in different positions; and the last group those conditions in which this latter was nearly always the case but the illusion occasionally present. The first column in each group gives the time relations, *i. e.*, the time of the first sound, of the interval and of the second sound, in sigma. The second gives the distance apart of the receivers and the next the distance from the receivers to the subject's head, in centimeters. The fourth column in the first group indicates the per cent. of trials yielding the impression of movement, this percentage being based on from 10 to 20 trials. About this same number of trials forms the basis for the conditions in the other columns. There were few cases in which a single sound was heard. In the visual experiments the two lines were often seen simultaneously. In the auditory field this simultaneity would naturally resolve into a single sound midway between the sources. This occasionally happened, usually with intervals of 10 sigma or less, and the timbre differences in the receivers militated against such results.

The table shows that, under certain conditions of time and distance, two similar sounds in different positions with a definite time interval between, give the impression of a single moving sound.

It is to be noted first of all that one of the subjects does not appear in the table. This subject never reported movement on any of the trials with two separate sounds. Her auditory acuity was very good, according to her own statement, and it was quite manifest in the present case. Even with a single receiver moving and the sound broken in the middle of the movement she was seldom deceived, apparently catching the click as the diaphragm made its first contact with the magnet. On the last hour's experiment, however, as a result of suggestion she reported the illusion frequently. The experimenter intentionally expressed surprise at the

TABLE I

Subject	Optimal Movement				No Movement			Seldom Movement		
	Time	Distance			Time	Distance		Time	Distance	
A.....	190-33-190	18	200	80%	150-100-150	27	50	190-33-190	27	200
	76-10-76	18	200	75%	225-25-225	18	200	175-30-175	9	200
	76-10-76	27	200	100%	120-70-120	27	50			
	76-8-76	18	200	50%	120-48-120	20	200			
	68-54-68	18	200	83%	114-15-114	9	200			
					105-85-105	18	200			
					80-70-90	27	50			
					72-28-72	20	200			
					60-60-60	20	200			
Bu....	90-10-90	18	200	66%	175-30-175	18	200	150-90-150	21	50
	67-6-67	18	200	50%	135-12-135	18	200	114-15-114	27	200
	47-38-47	18	200	66%	105-18-105	18	200	47-68-47	27	200
	47-38-47	27	200	50%	100-60-100	18	50			
	47-32-47	18	200	75%	79-55-79	18	50			
	47-32-47	27	200	75%						
	37-30-37	18	200	50%						
	37-30-37	27	200	66%						
C.....	225-25-225	27	200	66%	120-100-120	18	50			
	190-25-190	27	200	100%	90-10-90	27	200			
	180-48-180	10	200	100%	80-70-80	18	50			
	150-30-150	9	200	100%	70-80-70	18	50			
	150-30-150	18	200	100%	68-54-68	27	200			
	150-30-150	27	200	75%	66-100-66	18	200			
	150-30-150	36	200	66%	50-33-50	18	200			
	144-54-144	50	200	100%	33-25-33	18	200			
	144-54-144	60	200	100%						
	144-28-144	50	200	60%						
	114-15-114	27	200	100%						
	114-15-114	18	200	100%						
	108-28-108	10	200	75%						
	76-10-76	27	200	100%						
	72-28-72	10	200	66%						
	72-28-72	20	200	62%						
	67-7-67	27	200	85%						
	47-38-47	18	200	86%						
P.....	190-25-190	18	200	75%	175-30-175	18	200	150-90-150	21	50
	135-12-135	18	200	66%	120-48-120	10	200	144-54-144	10	200
	114-15-114	36	200	100%	100-70-100	18	50	144-54-144	20	200
	76-10-76	18	200	71%	79-55-79	·18	50	144-54-144	30	200
	76-10-76	27	200	50%	76-16-76	18	200	144-54-144	40	200
	68-12-68	18	200	50%	58-68-58	9	200	144-36-144	10	200
					58-68-58	18	200	144-36-144	20	200
					58-68-58	27	200	144-36-144	40	200
								105-18-105	18	200
								90-8-90	18	200
								90-8-90	27	200
								72-28-72	10	200
								60-48-60	10	200
								67-6-67	18	200
								38-28-38	10	200

fact she never noted the actual motion of the sound, whereupon she did note it repeatedly.[1]

There are further individual differences evident. *C* manifests the illusion to a far greater extent than any of the others, reporting it not only under more settings but also in a larger per cent. of trials on a given setting. In some instances the illusion was successful with the receivers 50 or 60 cm. apart. He stated that it was 'usually easy to tell which way they go and to tell one from two.'

P's results, on the other hand, fall mostly in the group in which the illusions occurred but seldom on a given setting. He often noted however a peculiar 'feeling of movement' in addition to the perception of the two discrete sounds. To quote his introspection: "Sometimes get two sounds but a distinct feeling of movement. I rationalize it with the stick moving and so get a feeling of movement." Or again: "Two sounds but an impression of movement besides. Get idea it is the same sound travelling. Have often some visual scheme, *e. g.*, like sticks on a fence. The feeling itself unanalyzable, like the feel of a band going down the street; sort of an interpretation of the thing." He noted at another time that, "If the sounds move at about the proper rate to coincide with what I think is moving I get the impression of movement."

A reported the illusion in only a few conditions and these were mainly near the outset of the experiment. Toward the end he became sceptical. He seemed to judge much by the after-image. The following introspections are typical: "Judgments afterward, like seeing a mass of things on a screen and analyzing it;" "Distinct image afterward and look at the sound picture." Toward the end of the experiments he admitted that he was sceptical and thought 'the apparatus fitted for only two sounds.'

Bu showed the illusion in a number of settings. He noted

[1] This effect of suggestion in producing the illusion indicates the advisability (with some subjects at least) of proceeding as in the present case without explaining the nature of the experiment, rather than following the methods used in the visual studies where the subject knew all about it and was required to state whether he perceived the illusion or not.

that it was easier to tell the direction in which the sound was moving than to tell one from two sounds. He found that much too depended on the attention. "If you wait too long it catches you on the wrong wave."

It is to be noted further that the illusion almost never occurred with the subject in the sound-cage, *i. e.*, 50 cm. from the receivers. Presumably the intensity of the stimuli was too great or their timbre differences too pronounced at that distance, even though the current was reduced as far as it was possible to do and still keep the fork going.

Wertheimer and Korte found that in general an inteıval of some 60 sigma between the two stimuli a few centimeters apart gave the optimal impression of visual movement. The above table would indicate a somewhat shorter interval as yielding optimal results in the auditory field,—ranging from approximately 10 to 50 sigma with the average 25 or 30. This may correspond to the usual finding that auditory reactions are quicker than visual.

The above investigators also found that the longer the exposure the relatively shorter is the optimal interval. A glance at the first column of the table shows that such was the case in the present study. The rows in the column are arranged for each subject in order of magnitude of the *exposure* with the longest at the top. It will be seen by following the *intervals* down the column, that in many cases they grow actually longer as the exposure decreases and in almost all cases relatively longer. There were a number of instances in which trials were made on a given subject with the exposure and interval in the same ratio while the absolute values were changed. Such series are grouped separately in Table II. The designations of the separate columns are the same as those of the first two divisions of Table I., but figures grouped together in a given row are those in which the ratio of interval to exposure time is constant. For example the first row indicates that with exposures of 76 sigma and an interval of 10, apparent movement was often perceived, while with exposures of 114 and interval of 15, which was in the same ratio as the first conditions, movement was never apparent. It is evident

that with three of the four subjects, given exposure and interval in the same ratio, the longer absolute values yield no movement while the shorter do. That is, the longer the exposure, the relatively shorter is the optimal interval.

TABLE II

Subject	Optimal Movement			No Movement		
	Time	Distance		Time	Distance	
A....	76-10-76	18	200	114-15-114	9	200
	76-10-76	27	200			
B....	67-7-67	18	200	135-12-135	18	200
	47-32-47	18	200	79-55-79	18	50
	47-32-47	27	200			
C....	225-25-225	27	200	90-10-90	27	200
	47-38-47	18	200	68-54-68	27	200
				33-25-33	18	200
P....	68-12-68	18	200	175-30-175	18	200

The other factor of interest in the present study was the effect of varied intensity. It was found in the visual experiments that, given the optimal conditions for apparent movement, increasing the intensity of the second stationary stimulus produced apparent movement in the reverse direction. This factor was studied in the present case for a few of the optimal time relations by adjusting the rheostats in series with the receivers. The sounds were equated in intensity at the beginning of the hour's work according to the judgment of the subject, and during the hour the rheostat in series with the second receiver was varied. The intensity of the sound was not measured, but four different settings of the rheostat were intermixed irregularly with the normal setting. Table III. gives the results found in this manner with two stationary receivers and the intensity of the second sound equal to or greater than that of the first. The first columns give the time of exposures and intervals and the distance of the receivers from one another and from the subject as in Table I. Then follow the per cent. of trials in which normal and reverse movement was produced in trials with the intensities equal and in trials with the second sound louder than the first.

Normal movement, of course, signifies apparent movement in the direction in which the temporal succession of the sounds actually occurred. The results in a given row in the table were always obtained on a single day.

TABLE III

				Equal Intensity		Second Sound Louder	
	Time	Distance		Normal Movement	Reverse Movement	Normal Movement	Reverse Movement
A............	190–33–190	18	200	60%	20%	0%	50%
	190–33–190	27	200	40%	0%	40%	20%
	76–10–76	18	200	75%	0%	27%	18%
	76–10–76	27	200	100%	0%	0%	0%
	76–8–76	18	200	50%	0%	60%	20%
B............	67–6–67	18	200	25%	25%	0%	66%
	47–38–47	18	200	50%	0%	0%	0%
	47–38–47	27	200	0%	66%	0%	100%
	47–32–47	18	200	50%	25%	0%	25%
C............	190–25–190	27	200	50%	50%	33%	66%
	114–15–114	27	200	66%	33%	43%	57%
	108–28–108	10	200	75%	12%	0%	60%
P............	135–12–135	18	200	66%	0%	66%	0%
	90–8–90	18	200	40%	0%	0%	0%
	67–6–67	18	200	33%	0%	0%	0%
Average.......	....	...	......	52.0%	15.4%	17.9%	32.1%

It is quite evident from the averages that the reversal effect is often produced by an increase in the intensity of the second stimulus. Whereas with stimuli of equal intensity the illusion of movement in the same direction as the actual temporal succession occurs in 52 per cent. of the trials and in the reverse direction in only 15 per cent., with the second stimulus louder the illusion in the normal direction occurs in only 17 per cent. of the trials and in the reverse direction in 32 per cent.

SUMMARY

1. The presentation of two faint similar auditory stimuli in quick succession a few centimeters apart yields, under certain conditions, an impression of a sound moving in the direction of the actual temporal succession of the stimuli. Individuals vary in their susceptibility to the illusion but four of the five subjects used manifested it at various times.

For the fifth subject imperfect technique combined with marked auditory acuity perhaps accounts for the results. A time interval between the sounds of 25 to 30 sigma yields on the whole the optimal impression of movement.

2. There appears a rather definite relation between the length of the period of exposure and of the interval between the stimuli. The longer the exposure, the relatively shorter must be the time interval to yield the optimal impression of movement.

3. If the intensity of the second stimulus is greater than that of the first the apparent movement is often in the reverse direction.

Conclusions

As far as the experiment was carried, the results were strikingly similar to those of Wertheimer and Korte. This fact necessitates a reconsideration of Wertheimer's theory. The fundamental assumption of this theory is that stimuli at various points in visual space are correlated with disturbances in corresponding regions of the visual cortex and that there is a 'physiologische Kurzschluss' between these cortical regions. Corresponding to this Kurzschluss there is the perception of (illusory) movement between the original external points. Increasing the intensity of the second stimulus can sometimes reverse the direction of this 'Kurzschluss.' The finding by Benussi of similar results in tactual space does no violation to the theory, for a similar assumption may be made of a correlation of different points on the skin with different points in the cortex and of a physiological 'Kurzschluss' between such cortical regions. But with audition it is a different matter. The auditory end organs are not stimulable at different points as are the retina and the epidermis. There is no evidence for the correlation of separate points in auditory space with separate regions of the sensory cortex. Wertheimer's theory cannot explain the results of the auditory experiments as it does those of the visual. Yet the same subjective phenomena are manifested in both cases. The implication is that there must be some additional factor operative in both audition and vision.

The writer is inclined to explain the phenomena on the basis of the 'action theory.' A sound at the level of the ears produces, by its binaural intensity difference, a motor impulse to turn the head or eyes in its direction. If a second sound in the same plane and at the same height supervenes shortly, there is a second motor impulse, presumably in the same centers, before the first is completely exhausted. The position of the sounds may thus be represented cortically by impulses of different intensity in the motor regions leading to the same muscles of the eyes or neck. If the second sound supervenes rapidly enough there is a continuity of the motor impulse. The direction in which the second impulse would lead, if executed, relative to the first, gives the cue as to the direction of motion. With the increase of intensity of the second stimulus it is conceivable that the second motor impulse is temporally facilitated sufficiently to produce the effect which would have obtained had it actually preceded. It is well known that reaction time decreases with the increase of the intensity of the stimulus.[1]

Some such factor as the above may well be involved in the similar visual and tactual phenomena. It seems probable that a visual or tactual stimulus produces an impulse to make some muscular adjustment in order to more clearly perceive the source of the stimulus. The biological importance of such a tendency is obvious. A second stimulus may produce an additional impulse before the first is exhausted and the continuity of the two gives the illusion of movement. At least it seems that a different theory from Wertheimer's is necessary to account for the auditory illusions. Perhaps this same theory may be involved in the visual and tactual inasmuch as they manifest the same characteristics. The writer hazards the belief that it is a question of the continuity of the motor impulses.

[1] Cf. Pieron, H., 'Recherches sur les lois de variation de temps de latence sensorielle en fonction des intensites excitatrices,' *L'Annee Psychologique*, 20, 1914, 17–96.

A COMPARISON OF DEAF AND HEARING CHILDREN IN VISUAL MEMORY FOR DIGITS

BY RUDOLF PINTNER AND DONALD G. PATERSON

Ohio State University *University of Kansas*

In the course of other individual tests with deaf children an attempt was made to measure the deaf child's immediate memory. One prerequisite for such a plan involves the presenting of relatively non-significant material in serial order to the subject who tries to reproduce the material in correct order and exactly as presented. This requirement was met by using digits arranged and presented serially, care being taken to avoid any commonly associated groupings such as historic dates or ordinary sequence, *i. e.*, 4, 5. Digits are relatively non-significant and as given in this series of experiments afford a fairly reliable index of the subjects' rote memory.

Four hundred and eighty-one deaf children were tested individually. They were all pupils attending the Ohio State School for the Deaf[1] during the school year 1914–15. The youngest pupil was seven years old, while the oldest pupil tested was twenty-six years of age. The smallest number tested in any one age was 15 (at age 7). The largest number was 51 (at age 16). All pupils of 19 years or more are grouped together and arbitrarily called adults.

The method used closely approximates that developed by Whitley.[2] A box whose inside measurements were $\frac{1}{2}$ by $1\frac{1}{4}$ by $15\frac{1}{4}$ inches was made with an aperture in the top measuring $\frac{3}{4}$ by $\frac{3}{4}$ inches. One end of the box was left open. Strips of

[1] The writers wish to acknowledge the kindness and coöperation shown them by the superintendent, Mr. J. W. Jones, the principal, Dr. Robert Patterson and the several teachers.

[2] Whitley, Mary T., 'An Empirical Study of Certain Tests for Individual Differences,' *Archives of Psychology*, No. 19, 1911, New York, The Science Press, p. 50.

cardboard were inserted in the open end of the box. The top card was then withdrawn slowly, exposing one digit at a time for about one second. There were two exposures of the same length of digits. Series of digits began with two digits and ended with seven. The score recorded was a plus or minus for each series of digits. Wrong order was counted minus. The digits were printed in blue type and measured $\frac{1}{4}$ inch in height. They were large enough to be seen clearly by every subject. After each series had been exposed, the subject immediately recorded in pencil on paper what he could remember. The series were always presented in the same order, beginning with the shortest and ending with the longest.

The memory span of each individual was determined by taking the highest series of digits reproduced correctly, irrespective of previous failures with fewer digits. This method approximates that used by Smedley.[1] By disregarding previous failures with fewer digits we are intentionally giving the advantage to the deaf child. His memory span as here recorded is probably a little higher than it would have been, had we used exactly the same method used by Smedley in determining the memory span of the hearing child.

The results are presented below in tabular and graphic form. A comparison is made between the oral pupil and the manual pupil, between the hearing child and the deaf child, between the deaf boy and the deaf girl, and between the congenitally and the adventitiously deaf.

Table I. and Graph I. show the comparison of the oral pupils with the manual pupils. The horizontal column at the top of Table I. gives the ages, the second column the numbers tested at each age, the third column the average memory span at each age, the next the span of 75 per cent. of the pupils, the next the span of the median pupil and the last horizontal column gives the span attained by 25 per cent. of the pupils. For example, there are 27 twelve-year-old orally taught deaf children. Their average memory span is 3.2;

[1] Smedley, F. W., Child Study Report No. 3 of the Department of Child Study and Pedagogic Investigation, Chicago Public Schools, 1900–1901, p. 53.

Table I

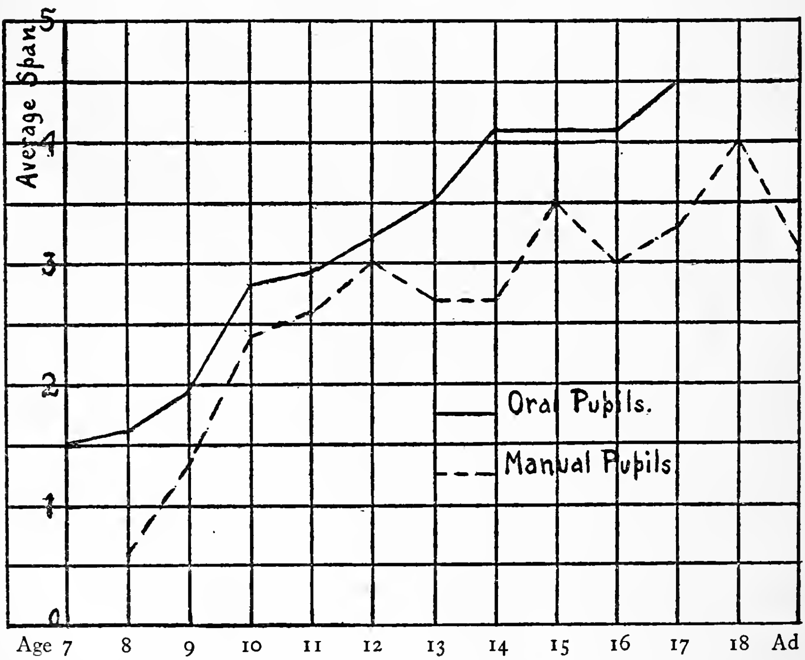
Comparison of the Orally Taught Pupils with the Manually Taught Pupils in Visual Memory

Oral Pupils

Age	7	8	9	10	11	12	13	14	15	16	17	18	Adults
Number	15	22	20	38	33	27	27	13	20	28	17	13	26
Average	1.5	1.6	1.9	2.8	2.9	3.2	3.5	4.2	4.2	4.2	4.5	4.5	4.5
75 percentile	0	0	0	2	2	2	3	3	4	3	4	4	4
Median	2	2	2	3	3	3	3	4	4	4	4	5	4
25 percentile	2	3	3	3	3	4	4	5	5	5	5	5	5

Manual Pupils

Age	7	8	9	10	11	12	13	14	15	16	17	18	Adults
Number		3	11	8	17	16	15	14	13	23	30	13	19
Average		0.6	1.3	2.4	2.6	3.0	2.7	2.7	3.5	3.0	3.3	4.0	3.1
75 percentile		0	0	2	2	2	2	2	3	3	2	3	3
Median		0	0	3	3	3	3	3	3	3	3	4	3
25 percentile		2	3	3	3	4	3	4	4	3	4	5	3



Graph I. Comparison of the Average Memory Span for Digits for Oral and Manual Pupils at Each Age.

75 per cent. of them have a span of at least two digits, the middle or median child has a span of 3 digits while 25 per cent. of them have a span of 4 digits or more. By referring to the lower half of the table we find that there are 16 twelve-year-old manually taught deaf children. Their average span is 3.0; 75 per cent. of them have a span of at least 2 digits, the median pupil has a span of 3 digits and 25 per cent. of them have a span of 4 digits or better. There is little difference between the two groups at this age. A comparison of the two groups, however, shows that this is not the case at other ages. Oral pupils, age 8 and 9, are superior to manual pupils of the same ages as shown by the medians. At ages 10, 11, 12, and 13 the medians show the two groups to be of equal ability. From age 14 on the medians show the oral pupils to be superior.

The averages for the two groups are shown in Graph I. The curves show clearly that at every age the oral pupils are superior to the manual pupils. These results are to be expected. Oral enthusiasts might be inclined to say that the oral method, which uses the teaching of speech and lip-reading as the sole means of instruction, is the causal factor in the superior memory ability of the oral pupils. This position is untenable because of a selective factor which operates to the advantage of the oral pupils. All pupils on entering the institution are instructed by means of the oral method. Those who fail to progress under this method are then relegated to the manual classes in which instruction is carried on by means of the manual alphabet and sign language. This failure to profit by the oral method is often due to inferior ability in other respects. In this test the inferiority is shown by a shorter memory span, on the average.

Table II. gives the data for all the deaf children and also the hearing children tested by Smedley and the deaf children tested by MacMillan and Bruner.[1] Graph II shows the relative abilities of the deaf children tested by us and the hearing children tested by Smedley.

[1] MacMillan, D. P. and Bruner, F. G., A Special Report of the Department of Child Study and Pedagogic Investigation on Children Attending the Public Day-Schools for the Deaf in Chicago. Chicago, 1908, pp. 70–71.

In the first part of Table II. we have given the actual distribution of the deaf children at each age. At age 7 for example 4 children had a memory span of 0, *i. e.*, less than 2; 10 had a span of 2 digits, and one a span of 3 digits. The total number tested at age 7 was 15, their average span 1.5, etc. The averages for the deaf children show a tendency to increase in memory ability from age to age. Those classed as adults have a lower average span than the 18-year-olds. This may mean that the 18-year-old pupils are on the average much better than would be expected in light of the scores of the ages below 18, or that the adult pupils are slightly poorer than they should be. Either factor or both may account for the results.

It will be seen that only 5 of the deaf pupils have a memory span of 7 digits. These are all oral pupils. Of the 16 deaf pupils who attained a score of 6 digits only 3 are manual

Table II

Memory Span of Deaf and Hearing Children

All Deaf Children

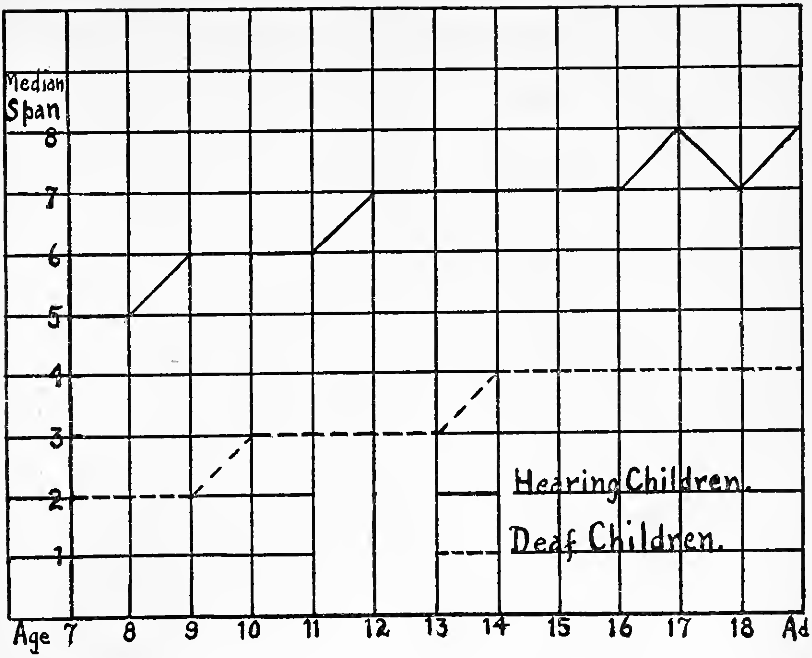
Age	7	8	9	10	11	12	13	14	15	16	17	18	Adults
0	4	10	12	2	1	1	1	2					
2	10	8	9	16	14	10	8	4	2	7	7		2
3	1	7	7	20	30	18	18	7	9	18	11	6	19
4			2	5	1	11	10	9	14	13	20	10	12
5			1	3	4	2	3	3	6	10	6	7	7
6						1	2	1	1	3	2	3	3
7								1	1		1		2
Number	15	25	31	46	50	43	42	27	33	51	47	26	45
Average	1.5	1.5	1.7	2.8	2.8	3.1	3.2	3.4	3.9	3.7	3.7	4.3	3.9
75 percentile	0	0	0	2	2	2	3	3	3	3	3	4	3
Median	2	2	2	3	3	3	3	4	4	4	4	4	4
25 percentile	2	3	3	3	3	4	4	4	5	5	4	5	5

Hearing Children (Smedley)

Age	7	8	9	10	11	12	13	14	15	16	17	18	19
Span	5	5	6	6	6	7	7	7	7	7	8	7	8

Chicago Deaf Children (MacMillan & Bruner)

Age	7	8	9	10	11	12	13	14	15	16	17	18	19
Number	10	13	13	17	12	7	12	12	17	6	4	2	2
Average span	3.3	3.3	3.7	4.1	4.4	4.7	4.8	5.2	4.9	5.7	5.5	5.0	5.0



GRAPH II. Comparison of the Median Memory Span of Deaf and Hearing Children.

pupils. Of the 52 pupils who have a memory span of 5 digits only 7 are manual pupils. The more efficient among the deaf pupils are predominantly the oral pupils.

In comparing the averages of the deaf children tested by us with those tested by MacMillan and Bruner we note that the Chicago Day-School deaf children are superior at every age. Several factors have combined to bring about this result. The Chicago deaf children reproduced the digits 'either orally or in writing.'[1] This is a variation in method which might greatly favor the Chicago children, since difficulty in writing possessed by any one child could be overcome by oral reproduction. It may be that there is a difference between results when reproduction is tested orally or written. It would seem to the writers that one method ought to be adhered to throughout. Further there was a

[1] MacMillan and Bruner, *ibid.*, p. 70.

high degree of selection among Chicago children tending to make them superior while the selection among the children tested by us was in the reverse direction. MacMillan and Bruner in their report (page 4) state that 10 per cent. of the deaf children in the day schools of Chicago were pronounced subnormal and excluded from the schools and sent to institutions in 1902 and further that since 1902 (page 73 of the report) "no feeble-minded children were permitted to enroll in the classes for the deaf." Such care in the selection of the deaf children in the day schools has without doubt appreciably raised the average ability of the children tested by MacMillan and Bruner. In Ohio there were five public day schools in as many cities during 1914–15. There was no doubt a tendency for the brighter deaf children to remain at home and attend these day schools while the duller ones from these same cities went to the state institution where they were maintained at public expense. Thus the selection works to the disadvantage of the state school. These are the factors which account for the evident superiority of the Chicago deaf children on this test. The comparison of the two groups of deaf children indicates the powerful influence that incidental factors such as variation in method and differences in selection of children may exert and hence warns us to analyze critically all the factors involved in comparative psychological studies.

Graph II. presents the curves representing the median memory spans of the deaf children tested by us and the hearing children tested by Smedley. It is rather startling to find that the average deaf child at any age never equals the average ability of seven-year-old hearing children. This difference is not shown on other mental tests. In learning ability the average retardation of these same deaf children on the Digit-Symbol test[1] was only three years. It is to be noted that both tests require the use of digits in their performance hence the marked inferiority of the deaf children in

[1] Pintner, R., and Paterson, D. G., 'A Class Test with Deaf Children,' *The Journal of Educational Psychology*, Vol. VI., No. 10, Dec., 1915, pp. 591–600. Also 'Learning Tests with Deaf Children,' *The Psychological Review Monograph Supplements*, Vol. XX., No. 4, Whole No. 88.

visual memory cannot be laid at the door of the materials used in this test.

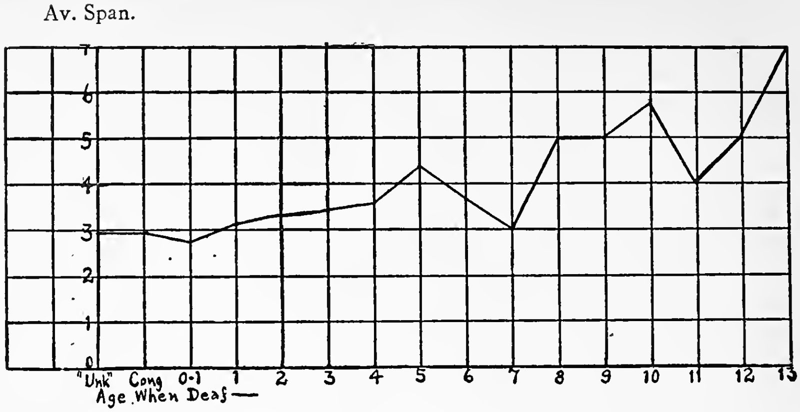
The cause of this disparity between the two groups must be sought in another direction. It is revealed by an analysis of the mental processes involved in this test. The hearing individual (in most cases probably) uses auditory images (consisting of the sound of the digits) plus inner tactual sensations aroused by the innervation of the muscles controlling the vocal chords, tongue and larynx. There may also be involved kinæsthetic imagery related to the hand movements necessary to write the digits. For the most part the auditory factor is eliminated for the deaf subject. This leaves the deaf child dependent for his memorization and recall mainly upon the visual percept, which becomes a visual image after the withdrawal of the stimulus. Many of the deaf children used their hands during the perceptual process, spelling out the digits as they were exposed. Many of them also used inner speech as indicated by lip movements. Hence memorization visually was in many cases aided by secondary sensory processes. Of course, a deficient nervous system caused either by poor heredity or the ravaging diseases which often cause deafness, probably accounts for much of the backwardness of the deaf in this test. But the results obtained in the Digit-Symbol test lead us to emphasize what is more probable, namely, the importance of audition in aiding the visual memory.

That the auditory factor is important is clearly brought out by Tables III. and IV. and Graphs III. and IV. Table

Table III

Average Memory Span of Deaf Children. Distribution according to Age at which Deafness Occurred

Age When Became Deaf...	Unknown	Cong.	0–1	1	2	3	4	5	6	7	8	9	10	11	12	13
Number..	52	206	32	62	55	27	14	13	6	2	1	3	4	1	2	2
Total score...	154	608	89	193	185	92	50	57	22	6	5	15	23	4	10	14
Average score...	2.96	2.95	2.78	3.11	3.36	3.41	3.57	4.38	3.66	3.0	5.0	5.0	5.75	4.0	5.0	7.0



Graph III. Average Memory Span of Deaf Children according to Age at which Deafness Occurred.

III. shows the grouping of the deaf children according to the age when they became deaf. The column headed 'unknown' includes all the cases whose records state that the time of the occurrence of deafness is unknown. That marked 'Cong.' includes those born deaf, that headed 0–1 those becoming deaf before they were a year old, etc. This table shows that the average efficiency tends to increase with the number of years of auditory experience prior to the deafness. The two deaf children who did not become deaf until they were 13 years of age have normal visual memory ability. There is a close correspondence in the first three columns. The 'unknown' cases do just about as well as the 'congenital' cases. This leads one to suspect that probably most of those cases recorded as 'unknown' were in reality born deaf. We seem to have experimental evidence which corroborates the statement of the superintendent of the Indiana Deaf School to the effect that probably two thirds of the cases recorded as 'unknown' are congenital deaf.[1] The conclusion to be drawn from a study of Table III. and Graph III. seems to be that in general the greater the previous auditory experience of the group the greater is the efficiency in immediate memory for digits.

[1] Johnson, R. O., 'Some Statistics Concerning Deafness,' Bulletin issued by the Indiana State School for the Deaf, p. 14.

Table IV. and Graph IV. show a comparison of the congenital deaf pupils with the adventitious pupils in visual memory for digits and in learning ability as tested by the Symbol-Digit test.[1] We find that the adventitious (cases in which the deafness is acquired or more precisely has developed after birth) are superior at every age, with the exception of age 9, to the congenitally deaf in memory for digits. This is not so when the learning test is considered. The conclusion to be drawn seems to be that in immediate memory previous auditory experience is of considerable importance as contrasted with another type of test that involves the same materials, *i. e.*, digits. It seems then that the startling backwardness of the deaf child in visual memory is in large part due to his lack of auditory experience.

Table IV

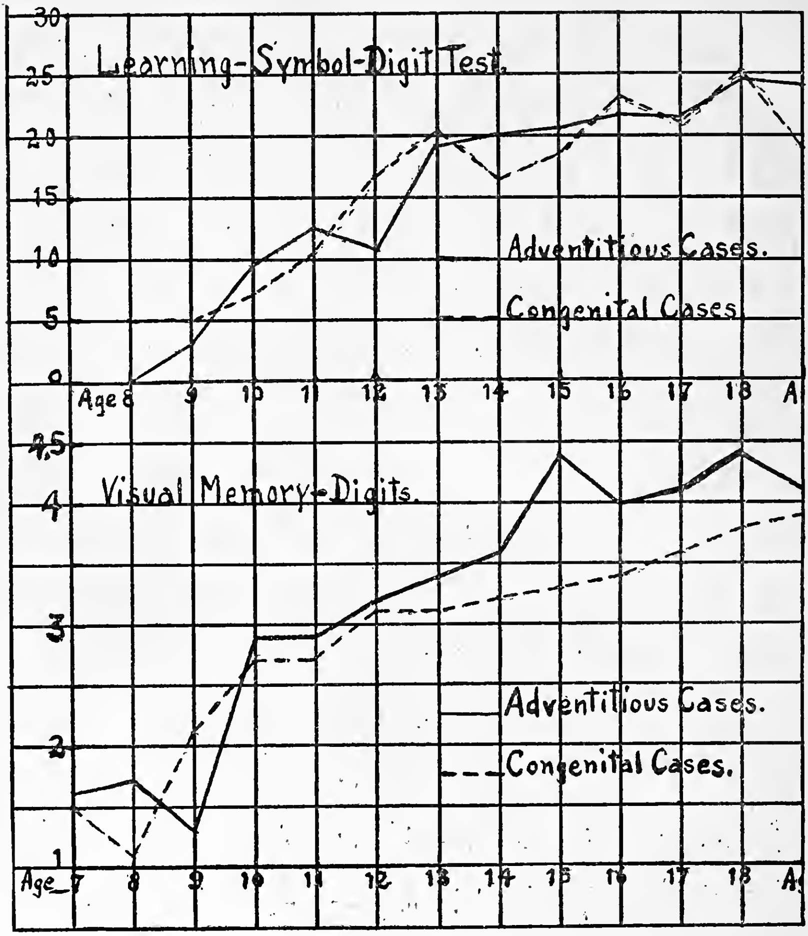
Comparison of the Congenitally and Adventitiously Deaf in Visual Memory for Digits and Learning Ability as Tested by the Symbol-Digit Test

Age	Visual Memory—Digits				Learning—Symbol-Digit Test			
	Congenital		Adventitious		Congenital		Adventitious	
	No.	Average	No.	Average	No.	Median	No.	Median
7	10	1.5	5	1.6				
8	13	1.1	12	1.7	5	5.0	4	0
9	15	2.1	12	1.3	8	5.0	8	3.3
10	25	2.7	20	2.9	22	7.4	15	9.8
11	22	2.7	24	2.9	22	10.6	21	12.2
12	21	3.1	19	3.2	16	16.9	22	10.7
13	15	3.1	19	3.4	11	20.4	22	19.3
14	11	3.2	13	3.6	16	16.7	17	20.0
15	14	3.3	13	4.4	15	18.8	24	20.5
16	23	3.4	22	4.0	24	23.4	19	21.8
17	15	3.6	25	4.1	10	20.9	28	21.0
18	9	3.8	17	4.4	11	25.4	21	24.6
Adults	14	3.9	28	4.1	14	19.0	32	24.0

Without attempting to enter into a lengthy discussion these results tend to discredit the widely accepted and prevalent theory concerning the vicarious functioning of localized portions of the nervous system. Here we have shown that the so-called visual center seems to have failed to take upon itself the functions of the auditory center and

[1] Results here given are taken from the Monograph previously referred to.

that compensation seems to be conspicuous because of its absence.



GRAPH IV. Comparison of Deaf in Visual Memory and Learning.

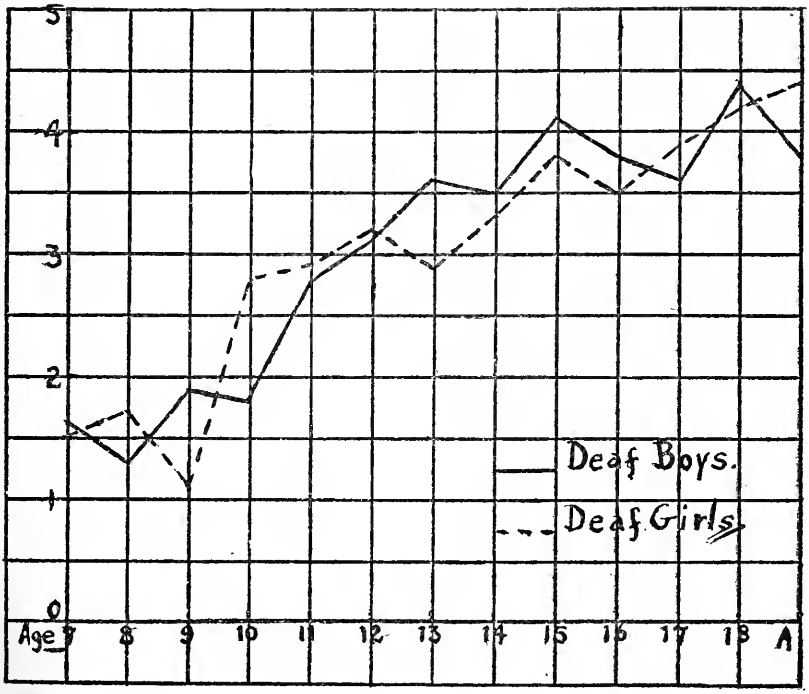
Are there any pronounced sex differences among the deaf pupils in visual memory? Table V. and Graph V. present the averaged results for the two sexes. Inspection of the table and graph shows that there is no constant sex difference. The curves of the two sexes cross each other repeatedly and one must conclude that a sex difference is lacking. Among

Table V

Comparison of Deaf Boys and Deaf Girls in Visual Memory for Digits

Age	Deaf Boys		Deaf Girls	
	No.	Average	No.	Average
7	7	1.6	8	1.5
8	14	1.3	11	1.7
9	23	1.9	8	1.1
10	23	1.8	23	2.8
11	33	2.8	18	2.9
12	28	3.1	14	3.2
13	18	3.6	24	2.9
14	13	3.5	14	3.3
15	15	4.1	18	3.8
16	31	3.8	20	3.5
17	22	3.6	25	3.9
18	11	4.4	15	4.2
Adults	21	3.8	17	4.4

Av. Span.



Graph V. Average Memory Span of Deaf Boys and Deaf Girls.

hearing children there seems to be a superiority of the girls,[1] but here as in the case of the learning tests there is no sex difference.

Summary

In summarizing the results of this test the following points are of importance.

1. Oral pupils are superior to manual pupils on the average.

2. Deaf children as a group have an abnormally poor memory span due to the lack of auditory experience.

3. Adventitiously deaf children are superior to congenitally deaf children on the average.

4. Auditory experience plays an important part in the efficiency of both hearing and deaf individuals in visual memory for digits.

5. Sex differences do not exist among the deaf in this test.

[1] Whipple, G. M., 'Manual of Mental and Physical Tests,' Part II., pp. 179–184, Baltimore, Warwick and York, 1915.

A NEW COMPLICATION APPARATUS

BY KNIGHT DUNLAP

This piece of apparatus was designed for use by a new method described below, requiring a large, clear dial. It was also deemed advisable to construct an instrument which would give ease and accuracy in setting a discrete stimulus in its proper place and such that the discrete stimulus would not vary in intensity with the rate of rotation. I believe that these problems have been satisfactorily solved in the present apparatus.

The apparatus consists essentially of a cast iron base, *A*, upon which are supported two pillars, B_1, B_2, 26 cm. high and 26 cm. apart. On the front of one of these pillars, B_1, is supported a dial, *C*, 36 cm. in diameter; on top of the pillars are brass bearings through which and through the center of the dial plate runs the main shaft, *E*. This shaft carries on the front of the dial an index hand, *F*—several hands of different length and shape are used. Just behind this support the shaft carries an arm, *G*, which can be clamped in any angular position with regard to the dial hand by means of a set screw and to which is attached a vernier scale, *H*, reading to degrees. By the use of the vernier the settings are made very easy, since the graduations on the circular scale are in ten degree units. Close to the other pillar, B_2, the shaft carries a large gear wheel, *I*, and a swinging arm, *J*, carrying a double gear, *K–L*, the smaller one of which is in mesh with the gear, *I*, on the shaft. This swinging arm can be locked in any position by means of a set screw in the pillar and passing through the slot in an L-shaped branch of the arm. The large gear on the arm is in mesh with the small gear on the shaft of the motor, *M*, fastened to the base of the instrument. By means of the swinging arm adjustment it is possible to use various sizes of gear on the motor shaft.

The adjustable arm attached to the vernier scale carries on its outer end a small horseshoe magnet, *N*. A vertical contact-lever, *O*, is carried on a support attached to the front pillar and in such position that the poles of the rotating horseshoe magnet, *N*, pass above and close to an iron crossbar on the top of the lever. A second horseshoe magnet, *D*, fixed on the support of the contact-lever keeps the contact closed until the traveling magnet, *N*, passes by and momentarily opens it. This contact device obviates the greater part of the noise which results from a mechanical contact breaker of the ordinary type and gives perfectly satisfactory contact. Since the vernier, *H*, and the index hand, *F*, are both adjustable on the shaft, the vernier may be set so that it reads zero when the index hand is actually at zero on the dial. Once set in this position, all further readings on the vernier will correspond to the readings on the index dial.

In order to secure a clean, single sound of intensity not varying with the rate of the interruption, an automatic relay[1] is employed in connection with a telephone receiver. The telephone receiver is mounted at the end of a wooden resonance chamber, somewhat like a megaphone, which greatly increases the sound. The primary circuit of the induction coil is interrupted by the contact breaking device on the complication apparatus. A secondary circuit which passes through the telephone receiver is interrupted on each side of the coil by two levers magnetically attracted by the core of the coil. When the current is flowing through the coil, the secondary circuit, *i. e.*, through the telephone receiver, is closed. At the moment of interrupting the primary circuit, therefore, the induced circuit flows through the telephone, producing a loud snap. Immediately thereafter, the core ceases to attract the levers and the circuit through the telephone receiver is broken on both sides of the core; consequently, when the primary circuit is reëstablished no secondary current is induced through the telephone circuit. In this way the break-current alone is used. By changing the

[1] This automatic relay was exhibited at the 1915 meeting of the American Psychological Association.

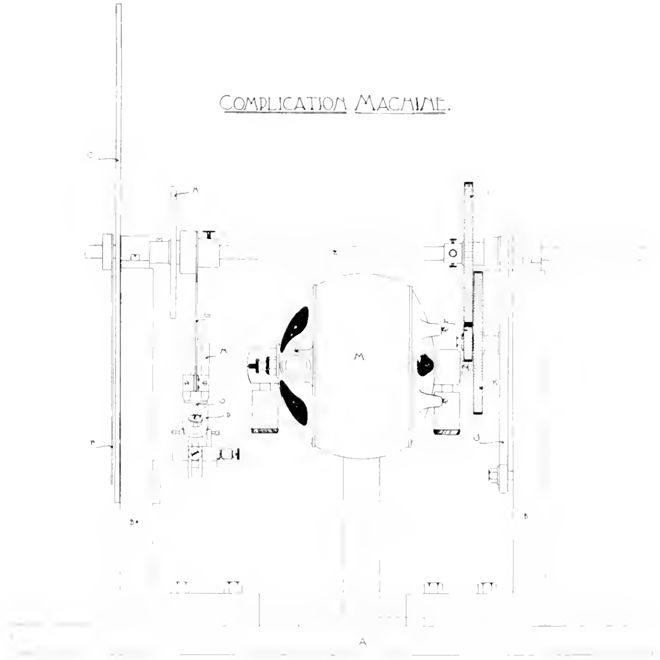


Fig. 1

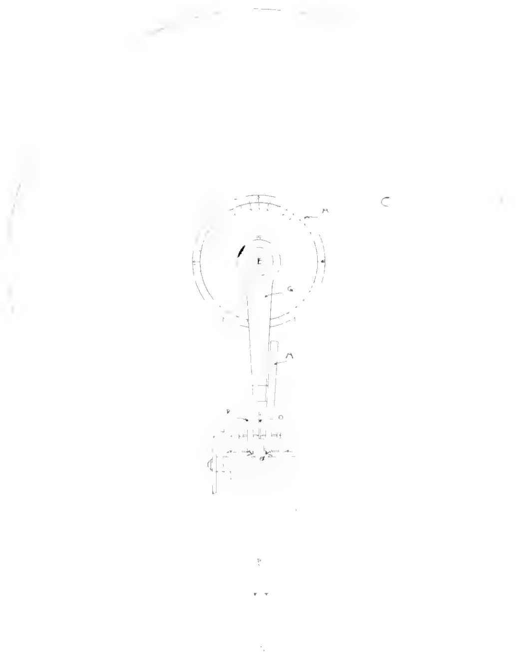


Fig. 2.

contacts on the automatic relay the make-current may be used, but is not so satisfactory on account of the variability. Since the current has the full time of the rotation of the shaft to establish itself, the break effect is almost absolutely uniform. Theoretically, there should be a slight difference in the loudness of the sound, depending on the rate of the interruption of the primary circuit; practically, no such difference is observable.

Another piece of apparatus which is used as an auxiliary to the complication instrument just described is a revolving shutter, composed of two sectored discs revolving in opposite directions on the same axis and in closely approximated planes. This shutter is placed in front of the eyes of the observer in such a way that the lines of vision of the two eyes are simultaneously cut off and then transmitted twice in each rotation of the discs. A metal plate with two eye-holes protects the observer from possible contact with the revolving discs.

On looking at the complication dial through the shutter revolving at high speed, the hand is not seen in motion but is seen in a series of different positions. By using an intense nitrogen lamp to illuminate the dial, the apparent brightness, which is reduced by the shutter, is brought up to normal. It would be possible to dispense with the shutter in front of the eyes and illuminate the dial by a beam of light interrupted by an episkotister, which would give somewhat the same effect, but for many purposes the revolving shutter is more satisfactory.

When using the revolving shutter, the observer is not only not tempted to make an eye reaction in the form of the following movement of the index hand, but finds judgment more difficult with such a movement. Hence, the avoidance of following movements is made very simple. This method does not, however, cut out all possible eye reactions.

DISCUSSION

In my article, The Influence of Inner Speech, etc., in the October 1916 number of this JOURNAL p. 365 ff., a few misprints occur. In the last line of Plate 7, the ditto marks under secs. and the words "R and" should be omitted. In Plate 8, " = 5" in the first line, and the ditto marks under mins. in the last two lines should be omitted.

H. B. REED

Journal of
Experimental Psychology

Vol. II, No. 2. April, 1917

SOME TESTS ON THE MEMORIZING OF MUSICAL THEMES

BY KATE GORDON

Carnegie Institute of Technology

A comparison is sometimes made by musicians between a musical composition and a logical or mathematical demonstration. The notes follow one another, it is said, in as strict and determinate an order as do the terms of a valid argument. One note implies the next. In brief, music has logical sequence, coherence and meaning. Now it is well known that the task of memorizing a given amount of material is greatly facilitated by the presence of logical meaning in that material. Hence it is reasonable to expect that if logical meaning exists in a given musical sequence, and is appreciated by the subject of the test, this fact will show in the memorizing process. The purpose of the following experiment is to find out whether music is memorized at a rate which allows it to rank with significant material or at a rate which classifies it with nonsense material.

Short musical selections were chosen, care being exercised to get simple singable intervals and a tempo which should be fairly uniform, and comparable with the series of nonsense syllables which were also to be used. The chief point, however, was to choose series which should give a unified musical impression. The following examples were used (Fig. 1).

The music was played by the experimenter on a piano, at a fixed rate of speed determined by the ticking of a clock. After each performance the subject was called upon to sing, hum or whistle as much as could be recalled. The selection was repeated until it could be rendered perfectly twice through by the subject.

FIG. 1

The accompanying curves (Fig. 2) give the distribution of twenty-five women, all but two being college seniors, and thirteen girls approximately eleven years old. The abscissæ show the number of repetitions necessary for learning the first selection, the ordinates give the number of individuals

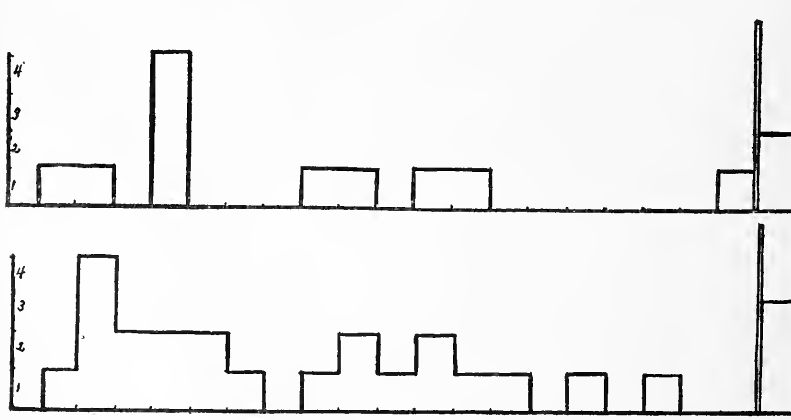


FIG. 2

learning it in a given number of repetitions. If the selection was not learned after twenty repetitions the subject was credited with a failure. Three adults and two children failed on this selection. Possibly these numbers are too small to be significant, but such as they are they tend to show a bimodal distribution both with adults and children, and they suggest that there is a fairly distinct break into a musical and an unmusical group in both cases.

In order to compare the learning of this musical example with the learning of nonsense material, a series of nonsense syllables was arranged having as many syllables as there are notes in the music. Since some of the notes are repeated in the music it seemed only fair to repeat some of the syllables also. This was done by assigning a given syllable to each note and repeating them as often as the notes were repeated. The series so constructed were:

I. non dil non rem bis yad non yad fim kel tek kel fim.
II. tob tob wab wab tob mel mel tob wab sil ven ven sil wab wab,
mel tob tob wab tob mel mel tob wab sil ven ven sil wab wab.
III. dez pum zif tov bax pum lar jed bax pum pum,
pum dez lib zif lib tov pum lar jed bax pum pum.
IV. mar zim lon mar ref zim mar bul ref lon mar ref.

These were read aloud by the experimenter, the speed being regulated as in the case of the music and being approximately

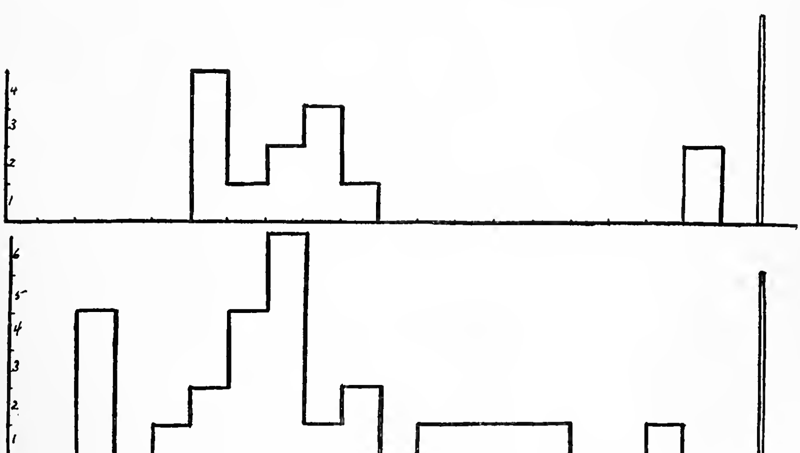


Fig. 3

the same. After each perusal the subject recited as much as could be recalled and so on until the series could be repeated twice without error. Fig. 3 shows the distribution of the same twenty-five women and thirteen children mentioned above for the first series of syllables. These curves approach more nearly to a normal uni-modal distribution. Now if we divide the adult group into two parts, on the basis of their facility in learning the first musical example we will get twelve whom we may call the musical group and thirteen the unmusical group. Then a comparison of the number of repetitions necessary to the learning of the music and the syllables shows that the nonsense syllables are easier for the unmusical group and harder for the musical group. The median of the repetitions for the musical group is 4, p.e. 1, for the music and 6.5, p.e. 2, for the syllables. For the unmusical group M. = 13, p.e. 3, for the music, and M. = 8, p.e. 1, for the syllables.

Now, although the musical subjects learned the music rather more quickly than they did the nonsense syllables, yet the difference is not nearly so great as the ordinary difference between nonsense material and significant material. I think this is due in part to the fact that the method of reproduction of the music was harder for the subjects. One of the most musical subjects said. "If I know the syllables I can say them, but I may know the music and not be able to sing or hum it." In order to avoid this difficulty and to get material which should be more nearly comparable than music and syllables the following plan was devised. The musical selections which had been learned in their normal form were reversed and played backwards. Thus Fig. 4 gives the

Fig. 4

first selection in Fig. 1 in reverse order, and it may be called a line of nonsense music. This line is comparable with its first form in every particular except logical sequence and finality: the number of notes, the pitch, the time, the intervals

are the same, and the rhythm is equally complex. But, since music is, in general, like language a non-reversible series, we may assume that the musical meaning has been eliminated. Of course it is sometimes true that a musical phrase played backwards happens, by chance, to give a musical impression, and to be musically good. But just so we must recognize that even language is occasionally reversible. For instance, in logic the universal negative proposition is always reversible, and certain cases of affirmative propositions happen also to be reversible. Tests, then, were made in which our subjects learned these reversed musical selections, and for the sake of further comparisons the nonsense syllables which had previously been learned in the order given above were re-learned by the subjects in reversed order. These tests were completed with twenty of the above twenty-five women. The number of repetitions was as shown in the table which follows. On the whole, then, and especially with the more musical subjects, the music is more easily learned forwards than backwards. The difficulty of learning the music backwards is, in all probability, greater than it appears in these figures, because in every case the practice effect favors the reversed musical series. The music had been learned in the regular forward direction before it occurred to the experimenter to use the reversed series, hence, instead of the practice being equally distributed between these two forms it is concentrated upon the second.

Musical Selection	No. of Subjects	Music				Syllables			
		Forwards		Backwards		Forwards		Backwards	
		Median	p. e.	Median	p. e.	Median	p. e.	Median	p. e.
No. I.......	Musical, 10	3.5	1.0	5.0	1.5	6.0	3.0	4.0	0.5
	Unmusical, 10	13.5	2.5	14.0	3.5	8.0	1.0	5.0	2.0
No. II......	Musical, 6	5.5	1.5	5.5	1.0	8.5	3.0	5.0	1.0
	Unmusical, 3	15.0	1.0	18.0	1.0	12.0	0.0	8.0	2.0
No. III.....	Musical, 4	5.0	1.0	13.0	1.5	10.0	2.5	5.0	1.0
	Unmusical, 2	8.5	1.5	13.5	1.5	13.0	3.0	6.0	1.0
No. IV......	Musical, 7	3.0	1.0	6.0	2.0	4.0	1.0	2.0	1.0
	Unmusical, 6	9.0	2.5	14.5	3.0	7.5	1.0	4.5	0.5

In the case of the nonsense syllables, on the contrary, the reversed series are learned more readily than the original ones. This is due in part, no doubt, to practice, but also to the fact that the syllables and groups of syllables retain their individuality and are recognized in the new order as the musical notes are not. The musical tones merge more completely into the new whole.

From the figures given and from the introspective comments of the subjects we seem warranted in making the following points:

1. For our musical group the musical selections were easier to learn than were the nonsense syllables. (A few tests were made of the recollection of music and syllables after a lapse of one to three weeks, with the general result that the music was better retained. In thirteen instances the musical selection could be sung after one presentation, but only in two cases was this true of the syllables.) For the unmusical group the music was harder to learn than were the syllables. The correlation between (*a*) rapidity in learning the music and (*b*) rapidity in learning the syllables is slight, if present at all. By the Pearson coefficient, for the twenty-five who tried the first selection $r = .043$, p.e. .134.

2. On the whole the difference between nonsense syllables and music, as indicated by the ease of memorizing, is less marked than the usual difference (*e. g.*, as given in Ebbinghaus) between significant material and nonsense material.

3. Mere familiarity with a musical selection seems to give it meaning. Both the subjects and the experimenter noticed that when the reversed music had been played over a great many times it came to "sound like a tune" and to seem more "reasonable."

4. Some striking instances occurred of the influence on memorizing of the perception of similarity. A glance at the second series of syllables shows that the last twelve of the two lines are the same. But, as the lines were read aloud and not seen, about half of the subjects failed to notice the degree of identity between these two parts. The subject would have learned correctly the first line and the first three or four

syllables of the second, and would then say that she did not know the rest. It took from one to six additional repetitions to learn the last part of the second line, when all the time the subject was actually reciting it as part of the first line. An analysis of the whole series and attention to the similarities would have lightened the labor of memorizing by a respectable percentage.

5. It is possible that a test of musical appreciation might be constructed which should hinge upon the difference in memorizing a significant and a non-significant musical selection. This would call for a set of themes which had been standardized both in difficulty and in musical quality.

KATE GORDON

QUANTITATIVE TONAL STIMULI WITHOUT QUALITATIVE CHANGE

BY HAROLD A. RICHMOND

Wesleyan University

The chief difficulty that confronts one in devising a method for producing quantitative tonal changes which may serve as stimuli for reactions is that of securing a change which is abrupt, but without click, or other qualitative modification. This difficulty is not lessened if one desires to use not only the beginning of the sound as a stimulus, but also its cessation, as well as increments and decrements in its intensity. For such experimental work on sound, Pillsbury[1] has pointed out with some care the superiority of the use of the telephone receiver over other methods. By changing the resistance in the telephone circuit, the intensity of the tone may be varied at will. Since, with the telephone receiver, "the intensity of its tone varies directly with the current," a measure of the current gives us the measurement of the tone intensity. The employment of the telephone for the cessation of a sound is objectionable because of the marked click which follows the breaking of the circuit, but Pillsbury showed that this difficulty can be obviated by superimposing two tones upon the telephone.

There are two disadvantages connected with the actuation of the telephone receiver as recommended by Pillsbury. First, unless both the tuning forks used in superimposing the tones are of a very high vibration frequency, several sigma must elapse before any change is complete. In the second place, the tone thus emitted by the telephone is not a simple tone, but is made up of a fundamental from the lower fork, and one or more overtones, from the higher toned fork and the natural period of the diaphragm. But, as Pillsbury states,

[1] Psychological Monographs, 1910, Vol. 13.

a simple tone would be more satisfactory than a complex tone or noise.

The method here presented makes possible the production of an abrupt, but smooth change in the intensity of a simple tone of high vibration frequency. As used in the Wesleyan laboratory, it has shown itself entirely satisfactory, both introspectively and in graphic records.

As a source of the required high frequency interruptions we use a small electric buzzer of high pitch, made by the Ericsson Mfg. Co., of Buffalo, N. Y. This is run by a single dry cell. A second circuit, also from a single dry cell, and flowing through the primary coil of an inductorium is interrupted by being connected across the contacts of the buzzer. Thus, although the second circuit is interrupted with the same frequency as the current actuating the buzzer, it is entirely independent. Hence, variations of resistance in the circuit through the inductorium will not modify the amount of current flowing through the buzzer magnets to change the rate of vibration.

The telephone receiver is actuated by the secondary coil of the inductorium, and thus gives a tone of the same pitch as the electric buzzer. The use of the inductorium is not a point at issue in the present discussion. In the circuit with the secondary coil and the telephone are various high resistance coils for changing the amount of current and the consequent intensity of the tone emitted by the telephone. For increment and decrement stimuli, some arrangement is necessary by which resistance may be put in or taken out at will. The accompanying diagram, Fig. 1, shows a satisfactory arrangement by which this is done. Included in the circuit represented by the unbroken line are shown the inductorium, the telephone receiver, and a constant amount of resistance, in this case 1,000 ohms. Two other resistances, labeled in the diagram 2,000 ohms and 10,000 ohms, respectively, are so arranged that one can be taken out of the circuit or the other put into it. This is accomplished by the simultaneous break and make key, *ABC*. In all cases the experiment started with a constant pre-stimulus tone con-

ditioned by both the 1,000 ohms and the 2,000 ohms resistance in the actuating circuit. By means of the make key the 2,000 ohms may be short-circuited to produce a louder tone, and by means of the break key the 10,000 ohms which has previously been short-circuited may be thrown into the circuit to produce a decrement in the tone.

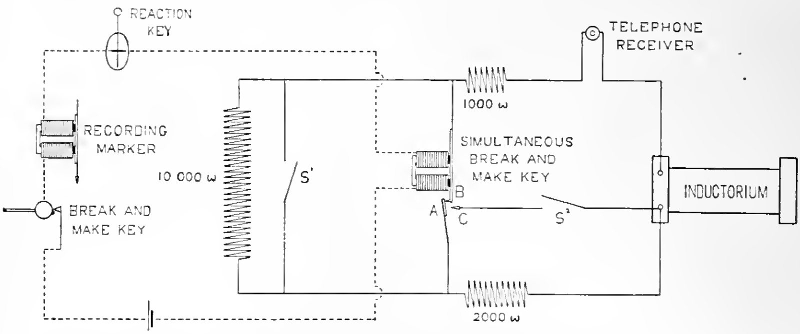


FIG. 1. Diagram showing plan of electric connections for giving stimulus and recording response in reactions to quantitative tonal stimuli without qualitative change.

The simultaneous break and make key is operated by an independent circuit from a storage battery, represented in the diagram by the broken line. In series in this operating circuit are also the subject's reaction key, a break and make key at the chronograph to determine the moment of stimulation and re-make the circuit, and the signal marker for recording both the moment of stimulation and the reaction. With the operating current flowing in the electromagnet of the simultaneous break and make key, the armature bearing the platinum contact point, *B*, is held by the magnet against the resistance of a spring. Contact point *A* is also mounted on a spring working in opposition to the pull of the magnet. As the armature moves away when released by the magnet, contact *AB* is broken at the moment when contact *AC* is made.

With the armature held against the magnet and with *A* and *B* in contact, as shown in the diagram, the switches, S^1 and S^2, may be open or closed. In either case the flow of current is through both the 1,000 ohms and the 2,000 ohms

resistance. When both switches are open, the break between *A* and *B*, upon the release of the armature, breaks the short-circuit across the 10,000 ohms resistance and throws this into series with the other resistances, thus producing a decrement in the intensity of vibration of the telephone receiver. When both switches are closed, the make between *A* and *C* short-circuits the 2,000 ohms resistance with a consequent increment in the intensity of the tone emitted by the telephone. Since both make and break in the actuating circuit occur at exactly the same time, the break in the operating circuit through the magnet coils which conditions the release of the armature and the consequent stimulus, may be used to record the moment of either increment or decrement stimulus. In our experiments, the break in the operating circuit was produced automatically by a successive break and make key at the chronograph, for which a falling plate device was used. During its fall, the plate engages the horizontal arm attached to the wheel contact of the break and make key shown at the extreme left in the diagram. As the wheel revolves, a small non-conducting offset on its surface pushes away the spring which forms the contact, thus breaking the circuit and causing a movement of the recording marker, together with the release of the armature of the simultaneous break and make key. The latter, by changing the resistance in the circuit through the telephone and the inductorium, gives the stimulus. Further revolution of the wheel carries the offset beyond the spring. The contact between spring and wheel is then re-made and the circuit again closed.

By means of its spring attachment, the released armature of the simultaneous break and make key is held so far away from the magnet that it cannot be drawn up again by the flow of current in the magnet coils. In re-setting, the armature must always be brought against the poles of the magnet by hand. After the breaking and re-making by the break and make key at the chronograph, the operating circuit remains closed until broken by the release of the subject's reaction key. It will be observed that the moment of the stimulus and the moment of response are recorded by breaks

in the circuit through the recording marker. By this familiar means, the latency of the recording marker is eliminated from the record.

In order to determine the character of the change in intensity of the tone emitted by the telephone receiver, photographic records of its vibrations at the time of change were taken. Fig. 2 is an enlarged reproduction of a record so obtained. The vibrations of the telephone diaphragm as shown in the lower half of the record have a magnification of approximately 6,000 times. In the upper half of the record is shown the vibrations of a marker run by a 100 v. d. tuning

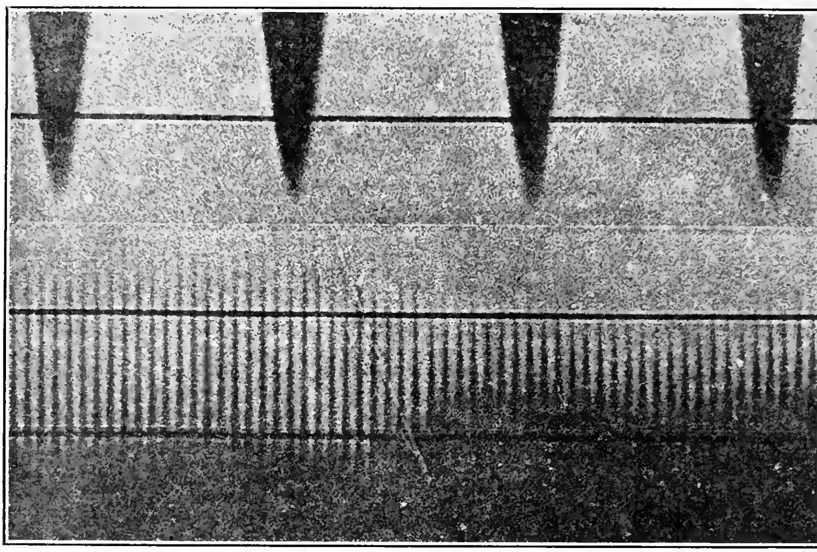


FIG. 2. Micro-photographic record of the telephone vibrations as the tone intensity is decreased. (Magnified 6,000 times.)

fork. Comparison with this shows the telephone receiver to have been vibrating at approximately 1,700 times a second, both before and after the change in resistance. The vibrations of larger amplitude at the beginning of the record were produced with a resistance of 1,000 ohms in the circuit. To change the intensity of vibration a resistance of 2,000 ohms was added, and the recorded change is consequently a decrement.

For photographing the vibrations of the telephone receiver, a small paper pointer was attached to the diaphragm. The end of the pointer was fastened to the middle of a fine silk fiber held vertically in front of the receiver by two supports. Thus the fiber was made to vibrate in exact correspondence with the vibrations of the diaphragm. These vibrations magnified by a projection microscope were photographed by a falling plate camera.[1]

Two important factors in the character of the change shown by the record should be noted. In the first place, the changed amplitude of vibration is attained in less than two full vibrations of the diaphragm. This means a time duration for the complete change of approximately a single sigma. In using a telephone receiver actuated by the commercial current of 60 cycles and a tuning fork of higher frequency, as was used by Pillsbury, even were the complete change accomplished at the end of a single full vibration, it would mean a time duration of over 16 sigma, during which there is a gradual change in the physical conditions of the tone; and since the change may not be completed in a single vibration, its duration is probably even greater.

The second factor to be noted is the smoothness of the change. There is not the slightest indication of a click, or any other form of disturbance attendant with the change. In the elapse of a single sigma, the larger amplitude of vibration has passed smoothly into the smaller. The change is purely a change in intensity.

Although the record here presented shows only the character of the change in a decrement in intensity, records taken of the increase change show it to have been of exactly the same character. Records were also taken of the movements of the diaphragm at a make and at a break in the actuating circuit. No click was audible either at the beginning or cessation of the tone emitted by the telephone. The records also showed no indications of a click, although in some instances certain minor irregularities were observable in the character of the waves.

[1] For a description of the falling plate camera see Dodge and Benedict, "The Psychological Effects of Alcohol," 1915, p. 79.

THE FOCAL VARIATOR.[1]

BY A. P. WEISS

Ohio State University

For many experiments in the psychology of the more complex mental states or reactions, it is often desirable to control the objective clearness of the visual stimulus so that the stages preceding the recognition of the stimulus may be studied.

The apparatus herein described and illustrated is essentially a system of lenses which are so related to each other that a visual stimulus may be projected on a ground glass screen in any degree of clearness from an unrecognizable blur, to clear definition or focus, and so that the degree of clearness can be accurately measured. A simple illustration will make its operation clearer:

Suppose it is desired to determine which of two specimens of handwriting is the more legible. The two specimens are placed side by side on a card in the holder 6 of the hood marked 1 in Fig. 1 of the plates.[2] This will result in an image of the specimens being projected on to the ground glass screen 5. If the lenses 2, 3, 4, are set so that the image will be completely out of focus it will be impossible to read either of the specimens projected on to the ground glass. If, now, the images are gradually brought into focus by shifting the lenses in the proper direction, a point will be reached at which it will be possible to read one of the specimens but not the other (provided they are not equally legible, but comparable in other respects). By continuing the clearing up process until the other specimen may also be read, the difference between the degrees of focal clearness at which both may be read can be used as a measure of the relative legibility of the

[1] From the psychological laboratory of the Ohio State Univ., Columbus, Ohio.

[2] This apparatus was constructed by A. P. Freund, mechanician at the university, from drawings furnished by the writer.

two specimens. In other words the specimen that can be read when it is out of focus most, is the most legible specimen. Where this seems necessary, the degree of objective clearness can be reduced to terms of optical dispersion or other absolute units. A simple method of calibrating the instrument will be described later.

The apparatus controls the following stimulus conditions:

1. The extent to which the stimulus card will be in or out of focus on the ground glass. The limits are an unrecognizable blur on the one hand, and a clear sharp image of the stimulus card on the other. Each of the intermediate degrees can be accurately reproduced as often as necessary.

2. The blurred image is always the same size as the clear image. This is of importance when the experimental conditions require that the area of retinal stimulation remain constant.

3. The image on the ground glass can be made either larger or smaller than the size of the stimulus.

4. The intensity of the illumination and the duration of the exposure of the stimulus or image may be varied within any limits.

5. The stimulus card holder can be removed and a tachistoscopic series or moving picture projection can be substituted.

The following are some of the problems which may be studied by the aid of the apparatus, either to secure new data or to check other investigations by a different method.

1. Relative legibility of handwriting, type faces and forms, diacritical marks, symbols, visual signals.

2. Higher thought processes,[1] association, recognition, attention, perception, illusion, imagination, expectation.

3. The relation between learning and the character of the visual stimulus as indicated by the question, "Of what importance is visual acuity in the learning process?" This problem is perhaps of greater importance in genetic and animal psychology.

4. In the analysis of the factors which make the different parts of a visual complex of unequal attentive or affective

[1] Dr. G. F. Arps, head of the department, is at present using the apparatus for an experiment in this field.

value, such as advertisements, pictures, geometrical figures, or esthetic combinations.

5. Analysis of individual differences as revealed by the stages preceding the recognition of some standardized stimulus card; and a comparison of either the introspective or verbal reaction series of different types of individuals.

Description and Method of Operation

Fig. 1 is an illustration of the focal variator set up ready for experimentation.[1] A cloth cover which stretches between the hood and the ground glass screen is not shown. This

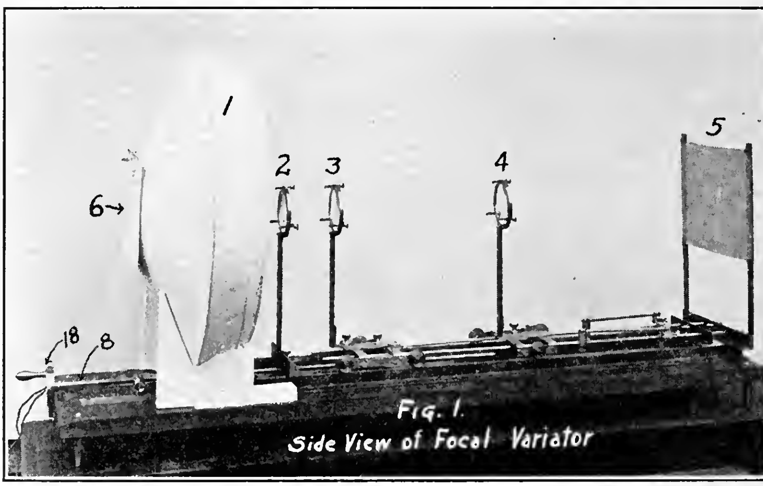


Fig. 1

cover hides the lenses from the observer and also prevents him from glancing around the screen to see the stimulus card. No. 1 (1) is a sheet iron hood, painted white on the inside. This hood contains eight electric light sockets distributed in a circle. The light from the lamps of these sockets is projected on to the stimulus card which is held in position by the trap door 6. Where it is desirable to expose the stimulus tachistoscopically, the door 6 is permitted to hang down and

[1] The number without the parenthesis refers to the same number on the plates. The number in the parenthesis indicates the number of the plate which is being described. Where no plate number is given the preceding plate number is implied.

the tachistoscope is moved to the proper position. It is not necessary that this adjustment be very accurately made since a much finer adjustment is possible by moving the lenses and ground glass. Where it is desirable to use lantern slides these may be placed in a special holder and substituted for the stimulus card, but it must be remembered that lantern slides must be illuminated from the back.

No. 2 is the stationary lens, double convex, focal length 12 inches (30.2 cm.) diameter 3 inches (7.1 cm.). This lens

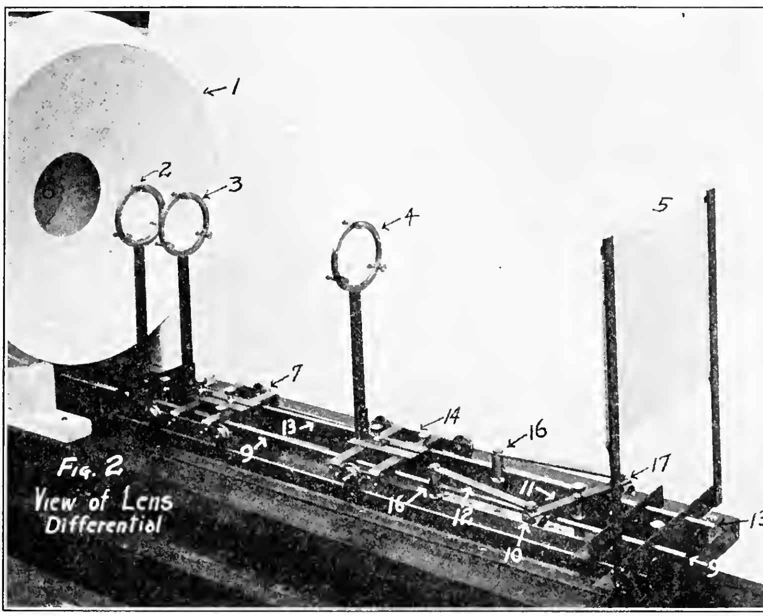


Fig. 2

is adjustable forward and back and controls the size of the image on the ground glass 5.

Nos. 3 and 4 (1). Differential lens system. Both these lenses have the same optic constants as No. 2. The variation in the clearness of the image is brought about by these lenses. They are interconnected in such a way that they travel in opposite directions. As one lens increases the size of the image the other reduces it and the final results of their effects is to produce a blurred image on the ground glass which,

however, is the same size as the clear image. The method of operation is as follows:

Lens 3 (2) is mounted on the three-wheeled carriage 7. This carriage is shifted back and forth by the screw 8 (1, 3). The rod 9 (2) is clamped to the carriage and moves with it. The clamp 16 fastens the rod to the pivot lever 11 by means of the connecting rod 12 and the clamp 10. A similar connection on the opposite side connects rod 13 also to the lever 11. This rod 13 is attached to the other lens carriage 14. If now lens carriage 7 is moved in one direction, its movement will be

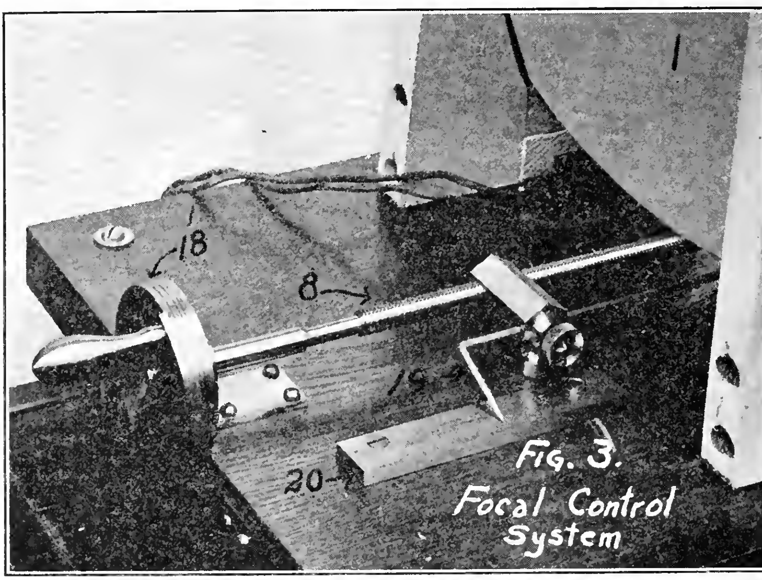


Fig. 3

transmitted to lens carriage 14 through the levers, rods and clamps, 9, 16, 12, 10, 11, 17, 16, 13 and finally to carriage 14. The lever 11 being pivoted in the center, the motion of the carriage 14 will be in the opposite direction to that of carriage 7. By shifting the lever connections 10 and 17 to various distances from the center of the lever 11, the system of lenses 3 and 4 can be moved with respect to each other at almost any kind of a rate. The clearing and blurring rate of the image on the ground glass is controlled by this device.

No. 5 (2) Ground glass screen. Size 8x10 inches (21.6x25.4 cm.). The image of the stimulus card is projected on to this screen through the lenses 2, 3, 4. The observer sits at a convenient distance for reading from the back of this screen. The clearness of the image which is projected will depend upon the relative positions of the lenses and which are under the control of the experimenter.

An "observation" may be made as follows:

The operator starts with the handwheel 18 (3) turned to such a position that there is no image on the ground glass. Then he gradually turns the handwheel so that the image gradually comes into focus. The observer is asked to report as soon as he has an "idea, association, meaning" or a reaction for whatever is on the ground glass. The "reading" is made by noting the position of the pointer 19 on the scale 20 and for fine work, the reading on the periphery of the handwheel 18. In this way the experimenter determines the position of the lens system at any point. This process is continued through the various "meanings" which the image may have for the observer until the image is at its clearest.

A second method of making an observation is to clear up the image in a number of equal steps by giving the handwheel a definite number of turns each time and having the observer introspect or report at each step.

By interposing a key or a make and break device into the light circuit, the stimulus can be exposed intermittently. The same result can be secured by interposing a camera shutter at one of the approximate focal points of one of the lenses.

The screw 8 (3) has 16 threads per inch (2.54 cm.) and each complete turn of the screw is indicated on the scale 20. The periphery of the hand wheel is divided into 100 parts so that it is possible to read to the one-hundredth of a sixteenth of an inch or to steps of .00625 inch (.0158 mm.). After all corrections for lost motion in clamps and nuts are made the readings can be made accurately to at least one-hundredth of an inch (.0254 mm.) and this is finer than will be necessary for ordinary purposes.

The readings of the instrument can be used in two ways:

1. Relative degree of clearness of the image. This is the simplest method. It does not give the absolute degree of blurredness and is to be used only where a *relative* comparison is adequate. In this case the screw readings are taken directly, as representative magnitudes.

2. Absolute degree of clearness, or the extent to which the focal point of the image will be displaced from the ground glass. The image will always be in focus *somewhere*. However, there is only one setting of the lenses at which the image is in focus on the *ground glass*. It is this principle which is used to calibrate the instrument. The degree of clearness or blurredness of the image is thus measured by the "inches that the image is displaced in front of the ground glass." When this displacement is zero, the image of the stimulus card is at its maximal clearness. All other displacements result in a blurred image whose size, however, is the same as that of the clear image.

In calibrating the instrument the clamps 16 (2) are loosened so that the lens carriages may be moved independently of each other. The three lens 2, 3, 4 are then adjusted so the image on the ground glass is in focus and of the proper size. The size is controlled mainly by lens 2. Next the pointer 19 (3) and hand wheel 18 are set to zero and all carriage clamps fastened.

For adjusting the size of the blurred image so that it will remain the same size as the clear image, the pin-hole camera principle is used. A round card with an eighth inch opening in the center is placed inside the guides of lens 4. This will eliminate the blur and always give a clear (though faint) image whether it is in focus on the ground glass or not. The next step is to adjust the clamps 10 and 17 on the pivot lever 11 so that turning the hand wheel to its extreme positions will not change the size of the image. When now the pin-hole screen is removed, the image should be in focus on the ground at only one setting of the hand wheel, but the blurred image at all other points will be the same size as the clear image.

In experimenting, the hand wheel is turned so that the focal point of the image moves forward (toward the hood)

from the ground glass. To determine the focal point of the image when this is not in the ground glass it is necessary to use a white cardboard screen in front of the ground glass. The distance of this cardboard from the ground glass gives the degree of blurredness of the ground glass image.

The process of calibrating the instrument is really simpler than appears from the description. No detailed knowledge of the laws of optics is necessary.

HIPP CHRONOSCOPE WITHOUT SPRINGS

BY HOWARD C. WARREN AND PRENTICE REEVES

Princeton University

In adopting the suggestion of Dr. Dunlap to operate the Hipp chronoscope without springs,[1] several schemes of wiring were tried. The problem was to secure an arrangement so that all switches and the Hipp could be operated by one experimenter. While not altogether ideal, the arrangement here described[2] has proved effective in the Princeton undergraduate laboratory and has been used in several graduate experiments.

The armature adjustments of Dunlap are used, and the spring counterbalanced by an added weight. The laboratory storage battery system, which is used, is kept well charged and a potential difference of 16 volts is maintained across the terminals of the magnets.

In order to have the working conditions identical in each magnet the same supply circuit is divided so that the upper and lower magnet circuits receive the same voltage. This is shown in the diagram (Fig. 1) where one lead from the storage is connected to one terminal of each magnet on the Hipp and the other lead to the upper and lower breaks.

A Harvard rocking key is used and is connected so that each side serves in a separate circuit. The Hipp is connected through the side 1, 2, 3 and the stimulus through the side 4, 5, 6. When starting a reaction the connections 1–2 and 4–5 are made. After starting the driving mechanism of the Hipp and warning the subject, the key is thrown so that both connections are broken. The connection 1–2 breaks the top circuit of the Hipp, the armature is pulled to the bottom

[1] Dunlap, K., "The Hipp Chronoscope without Armature Springs," *Brit. J. of Psychol.*, 1912, 5, 1–7.

[2] The authors are indebted to C. C. Brigham for his valuable assistance in this work.

magnets and the dial hands are set in motion. The top circuit is again reestablished by making the connection 2–3, as 3 and 1 are previously connected. The subject stops the dial hands by reacting to the stimulus by pressing a break key in the lower circuit. The time between breaking 1–2 and making 2–3 in the ordinary operating of the rocking key has proved satisfactory. On the other side of the rocking key the breaking of 4–5 releases the relay armature from the magnets, the spring pulls the armature to *M* and makes the stimulus circuit. In demonstration experiments in the introductory

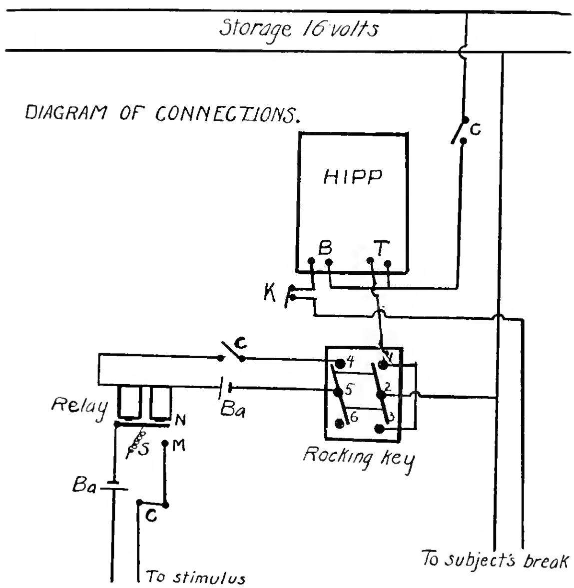


FIG. 1

course the spring-operated armature is considered sufficient, though a more accurate result may be obtained by having the armature operating between pairs of magnets. In this arrangement extra wiring is needed so that the same current may be used in both relay magnets. By making the distance *MN* about 1 mm., the stimulus may be considered as appearing when the clock starts, as the clicks of the Hipp and relay are simultaneous so far as observation reveals.

After the subject reacts, the driving mechanism of the

Hipp is stopped and the rocking key again set so that 1–2 and 4–5 are made. When 2–3 is broken in this resetting the Hipp armature goes to the bottom circuit, hence the extra key *K* is put in the bottom circuit to restore the armature to the top circuit before starting another reaction. Cut-out switches *C* are placed in the various circuits so that the latter may be opened during any pause in the experiment. All keys and switches are placed near the experimenter so that he may operate them without moving from the Hipp.

This apparatus is used for simple reaction to sound and light and for an association reaction to visual stimuli. For auditory reactions a bell is used as a stimulus, for visual reactions a small hole in a black box is illuminated by a light turned on by the relay, and in the association reaction the relay turns on a light behind thin paper on which the stimulus word is printed. In the last experiment the subject speaks the response word against the diaphragm of a Schall-Schlüssel wired across the "subject's break."

The subject is seated in one room and the experimenter in another with two closed doors between. A warning clicker in the subject's room is operated by the experimenter and communication carried on by prearranged signals. The relay and all noises are removed from the subject's room, so that the visual reaction is wholly a reaction to the visual stimulus.

SYSTOLIC BLOOD PRESSURE SYMPTOMS OF DECEPTION

BY WILLIAM M. MARSTON

1. Introduction.

The investigation on the galvanic effects of hidden ideas which have proved of interest both from a psychological and from a legal point of view have turned attention to the general problem of the physiological facts of the mental attitude in deceiving. The special problem suggested to me in the Harvard psychological laboratory was an investigation of the changes in blood pressure resulting from an effort to hide the truth. Just after we had begun the work, Benussi reported an experiment which in a parallel way studied the effect of lying on the changes in respiration.[1] He found a characteristic ratio of inspiration to expiration symptomatic of what he calls "internal excitement" caused by lying, and furthermore found this internal excitement to be much stronger in clever liars than in those easily detected, while, in the case of the latter, such excitement often tinged and modified the truthful records. Benussi, however, did not attempt to explain or analyze this "internal excitement," and his work leaves us with several troublesome questions in the answering of which this and similar methods of investigation might be invalidated. Is the "lying complex" sufficiently uniform in different individuals to be experimented upon as a unit? Through what physiological mechanism are symptomatic bodily changes effected? What, psychologically, is the nature of this "internal excitement"? Until these questions are at least partially answered we have in hand only a sort of psychological patent medicine, the ingredients of which, being unknown, may work as well one way as another under new conditions. Since, however, no definite emotional tests have

[1] *Archiv fur die Gesampte Psychologie*, 1914, pp. 244–271.

yet been established, a method similar to Benussi's seems inevitable in opening up a very complicated field. Benussi's results, indicating as they do great definiteness of lying symptoms, are sufficient to warrant the assumption of the uniformity of the deceptive consciousness as a working hypothesis. It will be the purpose of the present paper, in reporting the results of research on effects of this deceptive consciousness upon systolic b. p.[1] to analyze the data with a view of determining the physiological and psychological mechanisms involved.

2. Chief Physiological Factors of B. P. and Possible Psychological Influences upon these Factors

The blood, starting in the left auricle of the heart, is forced by the successive contractions of both auricles and ventricles into the aorta, or arterial stem. Thence the squeezing of the heart muscle forces the blood through the smaller arteries, and finally through the arterioles and capillaries into the veins, whence it returns under constantly diminishing pressure to the heart. In order to study psychic influences upon the b. p. it is necessary to bear clearly in mind the normal pressure conditions throughout the closed blood circuit. The pressure in the aorta is, of course, highest, and the broad channel of this artery offers comparatively little resistance to the blood flow; but as the smaller arteries are reached the factor of the friction with the arterial walls becomes more and more manifest, the side pressure and velocity pressure become less and less, and by the time the capillaries are reached the wall resistance is the dominating factor. The four chief factors in determining the arterial pressure at any given time may be said to be: (1) Heart-beat. It will be noted that the *rate* of the beat, and the *force* of the beat are two distinct functions; the former often increasing in inverse proportion to the b. p. while the pressure always increases in direct proportion to the force of the beat. (2) Constriction of the arteries, and especially of the arterioles and capillaries, usually called "peripheral resistance." (3)

[1] Abbreviation "b. p." will hereafter be used for "blood pressure."

Changes in elasticity of arterial walls. (4) Loss of blood. Since the last two factors are caused only by disease or by contingencies impossible of occurrence during the time occupied in taking any b. p. record, we need not consider them here.

Both the increase of heart beat and the increase of peripheral resistance, however, are factors to which we must look in accounting for any centrally caused changes. The systolic b. p. is peculiarly symptomatic of change in heart beat, while the diastolic becomes the crucial criterion if the changes we wish to study are brought about through changes in peripheral resistance.[1] Should both of these factors prove of essential value, then we must determine the pulse pressure, or ratio between systolic and diastolic. In order, then, to fix upon that aspect of arterial pressure which is the true indicator of that psychological complex which we wish to investigate, we must first glance briefly at the innervation of heart and capillaries, as well as at the psychological elements which have been found to substantially effect their functioning.

The rate and force of the heart beat are the algebraic sum of the cardio-inhibitory and accelerator nerve fibers. Thus a severing of the vagus nerves, or an inhibition at the cardio-inhibitory center will increase the heart beat as strongly as will a strong stimulation of the accelerator nerves. The latter seem to contain two groups of fibers capable of independent functioning, one group increasing the rate of beat, and the other increasing the force of the beat. It is only, however, through a reciprocal b. p. mechanism that the two groups function separately, central impulses seeming to set both groups in action. In the same way the peripheral resistance may be said to be the algebraic sum of the vaso-constrictor, and vaso-dilator nerves. Here, however, we must not think of peripheral resistance as a single organ or unit, but must remember, for instance, that the capillaries of the splanchnic area may be contracted while those of the skeletal muscles may be dilated. While, then, it is true that a strong vaso-constriction of an important area will immediately modify

[1] See *Boston Med. and Surgical Jour.*, Vol. CLXXII., No. 14, p. 530.

the diastolic b. p. of any artery above that area in the blood-circuit, it will only be a very general vaso-constriction which will have a significant effect upon the systolic pressure.

Since Professor Cannon recently definitely correlated six emotions with branches of the autonomic nervous system,[1] it will be well to further note the effects of these branches upon mechanisms regulating the heart beat and vaso-motor systems. The cranial division of the autonomic stimulates the cardio-inhibitory nerves, and the vaso-dilators of the stomach and digestive organs. The sacral division does not, in itself, innervate the heart, but only stimulates the vaso-dilators of the external genital organs and regulates the functioning of the excretory organs. When the excitement of the sacral system becomes sufficiently intense, the impulse passes over into the sympathetic. The thoracico-lumbar, or sympathetic, acts most uncompromisingly upon the accelerator nerves of the heart, and secondarily it inhibits the action of the digestive organs, contracting the blood vessels of these organs, and thus driving the blood to the skeletal muscles and outer parts of the body. The adrenalin released by the sympathetic impulses accentuates and prolongs this effect.

What emotions, then, will express themselves by heart acceleration, and which by vaso-constrictions? The answer is far from definite. The point first to be noted, however, is the uncertainty of any emotional influence through the vaso-motor apparatus upon b. p. The mild appetitive emotion is registered in the cranial division of the autonomic, and, consequently, has a vaso-dilator effect which, with the cardio-inhibitory action of this division, would be expected to diminish b. p. Yet, through a peculiar inhibition of the cardio-inhibitory center, a slight increase of pressure actually occurs. In the same way sex-emotion and relief, expressing themselves through the sacral division, would seem to tend to lower diastolic pressure through vaso-dilations, yet early in the development of sex excitement the sympathetic is aroused, and, until the climax is nearly reached, the effect

[1] W. B. Cannon, "Interrelations of Emotions," *Am. Jour. Psy.*, Vol. XXV., pp. 256–282.

upon systolic b. p. is scarcely discernible. Again, pain, according to Cannon, is one of the three major emotions normally expressing itself through the sympathetic division, and should, therefore, both increase the heart-beat and contract the blood-vessels in large visceral areas. There is no reason to doubt that such vaso-contractions occur, yet Binet early reported,[1] and his report is confirmed by the vivisectionists, that only the diastolic pressure is significantly altered by pain, that the heart is slowed in rate, and if there is any increase in systolic pressure it apparently is produced by the compensatory b. p. mechanism which operates to increase the force of the beat when the rate is diminished. Thus far, then, we have seen that the expression of emotion in vaso-motor modifications has little or no significance in determining the b. p. which would seem to be much more strongly and significantly controlled by the heart under normal conditions.[2] The one striking exception to this general rule forms the final argument for the choice of the systolic in testing the deceptive consciousness. Binet found that intellectual work at high concentration increased the diastolic b. p. 20, 30, and even 40 mm. The explanation is clear when considered in a teleological light. All the blood is required by the brain, and consequently, through vaso-constrictions, it is driven away from almost all other parts of the body. Yet, as I will indicate a little later, the systolic b. p. is not increased, nor is it significantly modified by even 40 or 50 minutes of mental concentration on study.[3]

The foregoing summary of the effects of minor affective elements and of intellectual work on systolic and diastolic b. p. will, I believe, justify the choice of the systolic in testing the deceptive consciousness. First, the use of the systolic eliminates the local effects of minor affective states; secondly, it eliminates the important and irrelevant factor of intellectual

[1] Binet et Vaschide, "Influence du travail intellectuel des Emotions, et du travail physique sur la pression du sang," *L'Anee Psy.*, 3, pp. 129–183.

[2] In medical diagnoses, of course, the peripheral resistance is often, on the other hand, the very condition to be investigated.

[3] Here we find the compensatory heart mechanism decreasing the force of the beat with the increase of rate.

work; thirdly, it is less susceptible to modification by physical pain than is the diastolic; and fourthly, it tends to record only the unequivocal changes in the b. p. system brought about through increase of heart-beat unimpeded by inhibitory reflexes or antagonistic functioning of the vaso-motor apparatus.

What mental processes may be expected, then, to cause an increase of heart-beat, and consequent rise of systolic pressure? We have, first, Cannon's three major emotions, rage, fear, and pain expressing themselves in the sympathetic division of the autonomic. Pain, as we have seen, has a less marked effect upon b. p. than the unobstructed operation of the accelerator nerves would seem to argue; yet it is possible that Professor Cannon had in mind not only physical pain, but also "psychic" or mental pain, which would resolve itself into an extreme degree of unpleasantness derivable from many sources. However, we may point to fear and rage as emotions which will, through the unobstructed operation of the sympathetic division, cause immediate rise of systolic pressure. It is well to note that only the smallest degree of fear or rage should, theoretically at least, be necessary to produce a rise in b. p. since the sympathetic system is the natural avenue for the expression of these emotions. Cannon also finds that sex-excitement, intense joy, intense sorrow, and intense disgust may, when they reach a sufficient level of intensity, break over into the sympathetic channels,[1] where they are felt merely as "excitement." Thus it appears that a profound modification of systolic b. p. cannot be analyzed with respect to its ultimate psychological causes, while any persistent smaller rise may presumably be attributed to rage or fear. Of course, very slight increases of pressure (especially if this be recorded by comparatively crude methods) cannot be regarded as significant of anything, but the necessary degree of intensity for emotions other than rage or fear to break into the sympathetic would seem to be so high that a considerable range of significant modification can be regarded as attributable almost exclusively to fear and

[1] W. B. Cannon, Ibid., p. 270.

rage. Although it is impossible to fix definite boundaries for this field, I shall hope to point out its general demarkations in considering experimental results. It only remains to point out, as a preliminary caution, the strong effect of physical exertion and of any contractions of skeletal muscles upon the accelerator nerves, and consequently upon the systolic b. p. All records must be keenly scanned, and conditions carefully controlled with a view to the elimination of the influence of this factor from the results. It is, however, much easier to control and allow for the factor of physical exertion, than it would be to exclude the element of mental work were we to use the diastolic or pulse pressures in examining the deceptive consciousness.

3. Method

The b. p. measurements were taken with a "Tycos" sphygmomanometer, an instrument substituting a spring for the mercury column of the older apparatus, and having the rubber pressure bag contained in a silk envelope made to wrap conveniently around the arm of the subject. The pressure was taken in the left brachial artery, the arm being completely bared before adjusting the instrument. This method of measuring the systolic pressure depends, of course, upon detecting the pulse in the radial artery, either by sphygmograph or by tactile sensations of the experimenter. Since the latter method was employed, the experiment is open to the criticism that the pulse is often present for a time after it has become impossible to detect it by mere touch. However, it may fairly be said that mechanical detectors have scarcely greater sensitivity, and are, in the long run, vastly less reliable. Moreover, the correction of this crudity of method would rather tend to accentuate increases in pressure than to diminish those found. Before starting the experiment, the experimenter practiced the taking of b. p. daily, for several weeks.

Four series were run off. The first three will be treated together, having been used upon the same group of ten subjects; and the last series, which employed the same method, will be introduced later, merely as a checking series.

Series A. (*Stories 1–8.*)—The subject came to the experiment as to an examination by a prosecuting attorney, resolved to save a friend who was accused of a crime. He sat down at a table beside the experimenter (but protected by a screen) and found on the table two papers face down; one marked "L" (Lie) and the other marked "T" (Truth). If, in saving his friend, the subject chose to lie, he turned over and read the "L" paper. This was a story prepared by the experimenter relating simple events, supposed to have been witnessed by the subject, and proving the friend guilty. At the end of the story were recorded certain facts, supposed to have been established by other witnesses, which the subject must admit in forging an alibi for his friend. He then proceeded, with these facts and the true story before him, to think out a consistent lying alibi. If the subject chose to tell the truth, he turned over the "T" paper, the contents of which were unknown to the experimenter, and found a consistent story, admitting the facts supposed to have been established, but completely exonerating his friend. This story was the *truth*, it was the only account he knew of the affair, and he told it as such. In either case the subject had 10 min., or until he announced he was ready, to thoroughly familiarize himself with the story he was about to tell, but was free to refer to the chosen paper any time he wished. The experimenter had prepared ten questions covering the incidents of the "L" story and an assistant had prepared the "T" story to successfully cover the questions, and the facts supposed to have been established. (After one "T" story has been told, of course, another was prepared by one of the assistants.) Thus it was impossible for the experimenter and jury to know whether the subject was telling a story of his own, or the one composed by the assistant. The questions were then put to the subject, and the jury closely observed his manner while answering. They then rendered a "verdict" as to whether he had lied or not, basing their judgment upon the internal consistency of the story as well as upon the subject's appearance while answering questions. These verdicts were written and passed in. The jury then left the

room, and the experimenter recorded his own judgment, which was based entirely upon the b. p. record. The subject gave his introspection, a final reading was taken and the instrument removed. The b. p. was recorded five times, in each experiment of Series *A*, (1) before the subject turned over the paper, (2) after he announced that he was ready, (3) after the fifth question, (4) after the last question, and (5) after his introspection. Complete notes of the subject's story were taken by the experimenter.

Series B. (*Stories 9–10.*)—The method of preparing the stories, etc., was exactly the same as in Series *A*. In Series *B*, however, the subject was first allowed to tell his own story without any questioning or interruption, and was then cross-examined by the experimenter or by the jury, or by both. If he chose he was permitted to reply to any question that "he did not know," or that "it wasn't on the paper." The jury then rendered their verdict orally, after whispered discussion, and the experimenter delivered the b. p. verdict orally. The jury then filed out, and the subject gave introspection, or in some cases, remained quiet for some moments. The essential innovation in this series was taking pressure readings every 2 min.

Series C. (*Experiments 11–12.*)—In this series the subject received an envelope with instructions sealed therein. He immediately left the room, and if he chose to lie he opened the envelope, obeyed the instructions, and came back and lied about what he had done. If he chose to tell the truth, he did whatever he liked for 10 min. and came back and gave an accurate and truthful account of his actions. After his return, he was given 2 to 6 min. to get his account in mind. In No. 11 no record of b. p. was taken until the subject returned, but in No. 12 a record was taken before he left the room. Owing to the lateness of the college year several subjects left before this last series could be completed.

Since Series *A*, *B*, and *C* can conveniently be considered together, I have tabulated together the results of these series for each subject.

The experiment was performed in the Harvard psycho-

logical laboratory, during the academic year 1914–15. Six subjects were graduate students of psychology, and four were undergraduates of considerable psychological training. The jury varied in number from 2 to 10, and was made up of men from Professor Münsterberg's elementary course. Beside this regular panel, several research students not in the experiment sat occasionally on the jury and numerous other students of psychology frequently acted as jurors. All who took part were greatly interested in the experiment, and all the subjects took the task of deceiving the jury very seriously, doing their utmost to outwit both jury and experimenter. The subjects were instructed at the beginning to choose an equal number of "T" and "L" stories, but although a list of previous choices was kept for each subject, this instruction could not be repeated without marring the conditions of the experiment. As a result, the subjects usually chose to lie more frequently than to tell the truth.

4. Results

A. Intellectual Work

In the above summary of physiological factors involved in systolic b. p. it will be remembered that the statement was there made that no one of such factors was influenced by intellectual work. This statement is substantiated by preliminary tests made upon all the subjects who took part in the experiment. B. p. records were taken while the subjects were doing arithmetical work at high concentration, while they were studying for college courses, and several extra short records were taken while inventing stories similar to the ones necessary for this experiment. A few significant variations were found, all of which could be directly correlated with some intense emotional intrusion; but during the actual mental concentration no uniform curve could be found either for all the men, or even for different records of one subject. No rise of more than 4 mm. was noted, and although the general tendency seemed to be a diminishing of pressure during a long period of mental work, no very significant or uniform descent of the curves could be noted. A single typical curve

will present the general results on this point as clearly as would an extended tabulation.

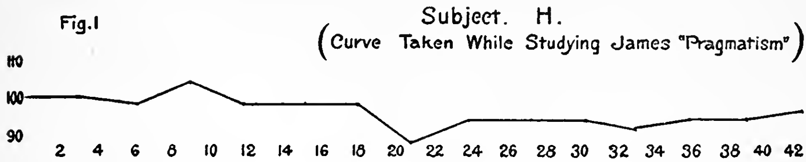


Fig. 1

It will be noted that this record extended over 42 min., in which time the subject covered 34 pages, with an excellent mastery of the thought therein contained, pressure being taken every three minutes. The major part of the curve is well below the initial pressure, and the single rise of 4 mm. is neither sufficiently high nor sufficiently prolonged to be significant. How then shall we account for such small, irregular fluctuations? It seems that we need not look beyond purely physiological causes. Besides the frequent minor irregularities of normal heart beat and vaso-motor adjustment, the factors of respiratory waves of b. p. and the longer waves due to rhythmical variations in the tonicity of the vaso-constrictor center under unusual conditions must be taken into account. In the light of the numerous and inevitable variations due to such constantly acting causes, it seems a safe general rule to regard no systolic variation below 6 mm., and perhaps none below 8 mm., as significant of major emotional influence. It is certainly possible that the intellectual work, besides raising the peripheral resistance and diastolic pressure, may also effect respiratory and minor chemical changes which cause systolic variations; but such physiological effects, at all events, seem to depend largely upon the temporary condition of the individual organism, and so may be dismissed as unimportant for the purposes of the present experiment. A careful study of the pressure changes during the "preparation" periods of the experiment shows a general result very similar to the preliminary intellectual work tests; while correlation of introspective reports of where intellectual work became necessary in the course of

the narrative with the pressure record at such points fails to show any increases of pressure.

B. General Results of Effect of Deceptive Consciousness

In a word, a uniform and significant systolic pressure curve was established by the results, as symptomatic of the deceptive consciousness. A rather surprising secondary result was the appearance of an almost equally definite Truth curve. Before proceeding to analyze the general aspects of the significance of these b. p. modifications, it may be well to note the two most general and plausible doubts which can be advanced concerning the possibilities of significant results under the conditions of this experiment. (1) Is the necessity for deceit sufficiently vital to furnish emotional stimuli for significant rises in b. p.? (2) If a sufficiently intense emotional situation is produced, will not the presence of witnesses, etc., cause exactly similar emotional influence while the subject is telling the truth? Let us glance, in answer, at the highest lying curve obtained in contrast with the lowest truth curve of the same subject, and at the lowest lying curve obtained contrasted with the highest truth curve of that subject.[1]

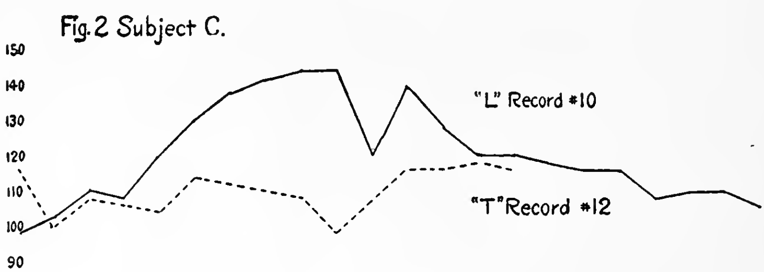


Fig. 2

The two *most* extreme curves, taken from the record of Subject *C*, conclusively indicate that neither doubt invalidated all the records taken, and the curves representing *least* extreme b. p. differences for truth and lying establish the significance of the pressure modifications in all the tests taken. In No. 10, Subject *C*, it can not be doubted that some very

[1] It would be obviously impossible to compare the lying curve of one subject with the truth curve of another, on account of the individual differences of initial and average pressure.

strong emotional situation was produced by the conditions of the experiment, since it will be noted that a total rise of 46 mm. in the b. p. was recorded. In No. 12, on the other hand, the central processes of the same subject, under the same conditions of witnesses, cross-examination, etc., caused a *drop* in the b. p. of 18 mm. toward the end of the cross-examination, with a return to initial pressure when the story was over, and a rise of only 2 mm. as the verdict was delivered. This and similar records seem to indicate very clearly that, during the telling of a truthful story to a suspicious and critical audience there is a more or less typical emotional (or other central) grouping of conscious factors which tend to inhibit any general emotional reactions to environment capable of increasing pressure, and which exert a positive influence over physiological conditions. How strong and consistent this influence may be can only be determined by a study of the individual results. The 8 mm. rise shown in record No. 11,

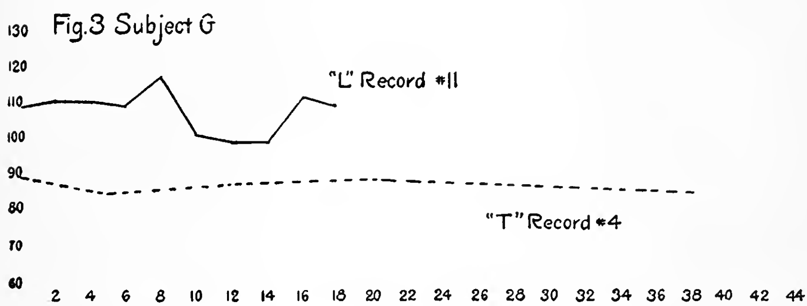


FIG. 3

Subject *G*, is the smallest b. p. increase recorded during deception by any subject and appears just on the edge of the field of pressure modifications significant of major emotional influence. It would seem that deception wrought little havoc in Subject *G*'s emotions, but when we glance at record No. 4, we see that while telling the truth, *G*'s b. p. did not rise *at all*, dropping 4 mm. in his highest truth record. The results, then, of this experiment unquestionably show significant b. p. changes under the influence of the deceptive consciousness.

But what constitutes the "significance" of a pressure curve symptomatic of deception? What differentiates such a curve from any chance rise of b. p. caused by the arousal of some incidental emotional complex? The answer to these questions is to be derived from a study of the above typical curves.

1. The amount of the rise is, in all "L" curves, too great to be accounted for by moderate degrees of intensity of any emotions other than fear or rage, the minimum rise being 8 mm.

2. The duration of the rise is, in "L" curves, too long to be symptomatic of a sudden and transient emotional association, the minimum duration of any rise being 8 minutes.

3. The rise of an "L" curve occurs in regular, climactic manner. The pressure starts its rise close to the beginning of the recital in every record as in the typical curves above, climbs with varying abruptness but with great consistency of movement to a definite climax, and then recedes. Subsequent questions may cause secondary climaxes, but these are patently subsidiary to the steady, persistent climb and fall of the pressure curve taken as a whole.

4. The apex of each curve is correlated very closely with that point in the subject's testimony which marks the crisis, or climax, of the whole "job" before the subject. This was determined partly by introspection, but chiefly by observations on the manner and attitude of the subject, and by noting the whole construction and plan of the false "alibi." Thus, like the other elements of "significance" in "L" curves, such correlation is capable of *objective* determination.

FIG. 4[1]

Subject..........	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Av. rise in "L" curve..	12.6	14	26.7	18.5	17.3	17.2	14	15	10.8	18.2
M. v..................	2.2	4	10.2	7.7	3.3	4.9	4	1	1.7	3.4
Highest rise in "T" curve.............	+6 and −10	+8	+18 and −4	+4 and −4	+6	+14	−4	−6	−10	−6

The table in Fig. 4 will present a rough summary of the

[1] Since closer fractional determination would be meaningless, the averages are given as of the nearest millimeter.

general results, the extremes of which have already been shown in Figs. 2 and 3.

It will be noted, from this table, that every subject's average b. p. increase during deception is well above his highest "T" curve, and that, with 2 exceptions, the highest "T" pressure mark plus the m. v. from the "L" increase is still below the average "L" rise. It is further true, although it does not appear in the above table, that, with the same two exceptions, every subject's lowest "L" curve was significantly higher than his highest "T" curve. The m. v. s. are, on the whole, low, seeming to suggest a rather fixed amount of b. p. increase for each subject, although the number of measurements is by no means sufficient to prove such a generalization.

A total average of 16 mm. rise in b. p. during 56 deceptions, by 10 different men, all such "L" curves having the significant characteristics pointed out above, seems conclusive proof of marked modification of b. p. during deception.

C. *Individual Results*

In order to determine, if possible, the psychological causes of the b. p. modifications during deception, as well as to study the uniformity of a possible truthful complex, it is advisable to review the individual records, subject by subject. Only in this way can the curiously close correlations of introspection and pressure record, the individual peculiarities, and certain interesting mixtures of truth and lying be considered. I shall attempt to bring out these salient features very briefly with each set of records, summarizing, thereafter, my own conclusions as to meaning and interpretation.

Subject A.—The stories composed by Subject *A* were, on the whole, very poor alibis. They were rambling, indefinite, rather wild, and very improbable, yet while telling the truth this subject managed to convey, by his peculiar manner of narration, the impression that he was lying, so that the jury found it very hard to judge correctly in any case. Subject *A* introspected during deception, a feeling of "responsibility" and *fear* of questions to come. He found lying "restful,"

Fig. 5[1]

Subject A

T					
2		4		11	
	120		116	2 2	118 118 118
End of preparation			←		←
	118	6	114	2 2 2	118 124 108
End of ques. 1–5			←	End of recital	
	120	11	118	2 2 2 2 2 2	118 116 114 116 108 108
End of ques. 6–10			←	End of cross-exam.	
	118	9	116	2	112
End of introspection			←	Verdict given	
10	120	10	116	2 2	114 118

[1] Number at head of each vertical column indicates which story was used in that record. Three digit numbers are mm. b. p.; and one or two digit numbers in narrow columns indicate min. s. elapsing between b. p. readings.

Fig. 5 (*Continued*)

Subject A

L												Conscious of Detection		T and L	
1		3		6		7		9		10		5		8	
	120		118		120		118	3 2 2	116 120 124 122	2 2	118 120 124		118		128
	←		←		←		←		←		←		←		←
	124		128	6	124	4	130	2 2 2	128 128 126	2 2 2 2 2 2 2	130 134 136 130 130 126 126	6	112	5	128
End of ques. 1–5			←		←		←	End of recital			←	End of ques. 1–5			←
	128		124	10	132	7	128	2 2 2 2 2	126 122 124 126 126	2 2 2	124 128 122	19	112	11	124
End of ques. 6–10			←		←		←	End of cross-exam.			←	End of ques. 6–10			←
	100		132	8	116	11	126	2 2	118 116	2	120	8	118	10 4	122 124[1]
End of introspection			←		←		←	Verdict given			←	End of questions by jury			←
	118		114	12	122	18	120	2 2	116 118	2 2	128 122	10	108	2	128
								End of introspection			←	Period of quiet			←
								2 2	116 116	5	118	10	108	11	126

[1] Lied

lax, and pleasant; but while telling the truth his feeling tone was "indifferent."

Notable Individual Records. No. 3.—At the 5th question subject felt "relief," and "elation" at supposed success in fooling jury. It will be noted that the pressure falls at this point, and rises during introspection when "*worry*" was felt lest he had not told a good lie after all.

No. 5.—Objectively, this story was a wild lie. Yet at the very first question, subject realized that he had betrayed himself to the jury and experimenter, and felt "disgust," "shame," and subsequent boredom. It will be noted that, unlike most of *A*'s "T" records, this consciousness produced a consistent drop in pressure. The most salient characteristic of the introspection was utter lack of interest and complete relaxation. No. 6 shows a similar drop from 132 to 116 in 8 min. after subject betrayed himself, consciously, in the 10th question.

No. 8.—Subject ran hard, just previous to coming into experiment, for about a quarter of a mile. The persistence of the influence of strenuous physical exertion is to be seen in the record, 26 min. being required for pressure to return to 122 (approximately normal). A single lie, told in answer to the last question, with introspective confession of this lie, sent the pressure up again 6 mm. in as many minutes.

Subject B.—*B* introspected, when lying, *fear* of many things, and it was for this reason that he did not choose to lie more often—he feared to fear! although, when it was over, he found he enjoyed the deception more than the truth. His stories were both good, although not ample, and his manner of telling both truth and falsehood was even and quiet. *B* felt "tense" during both deception and truth, the lying itself being more pleasant, and keeping him more "alert." It will be noted that a majority of *B*'s "T" records show consistent downward tendency, with a return toward the initial level.

Notable Individual Records. No. 8.—It will be noted that a severe pain raised the whole level of the day's blood pressure much above the subject's average level, and that this influence

evidently counteracted the usual downward tendency of *B*'s "T" curves.

No. 9.—The pressure was probably sent up, during the recital, by odd facial expressions involving considerable

FIG. 6

Subject B

T										L			
4		6		8[1]		9		11		5		7	
	96		90		106	3 2 2 2 2	84 86 86 88 88 86	2 2	102 106 94		88		*120*
End of preparation			←		←		←		←		←		←
21	84	10	84	3	106	2 2 2 2	86 90 84 92	2 2	88 88	9	96	4	*138*
End of ques. 1–5			←		←	End of recital			←	End of ques. 1–5			←
6	88	6	92	9	108	2	86	2 2	92 94	9	98	7	128
End of ques. 6–10			←		←	End of cross-exam.			←	End of ques. 6–10			←
8	88	6	92	8	104	2	86	2	96	7	92	14	112
End of introspection			←		←	Verdict given			←	End of in-trospection			←
8	92	10	96	10	108	2 2 2 2	88 100[2] 88 90	2 2 2	98 94 98	11	90	10	110
						End of introspection							
						2 2 2 2	90 90 88 88						

[1] B has *severe toothache.*

[2] Muscular contractions.

muscular contractions. A contraction of the left bicep, just as the pressure was taken, accounts for the abrupt rise to 100 mm. after the verdict was given.

No. 11.—A little physical exercise probably raised the initial pressure 2 to 4 mm.

No. 7.—Introspection revealed that for some minutes before coming into the experiment, *B* had been planning to deceive the jury. This would seem to have raised the initial pressure well above normal level (*B* had done no physical exercise), the actual lying sent it still higher, and it was still on the downward trend toward normal level when *B* left.

Subject C.—This subject did a great deal of laughing, but aside from this made an excellent witness, telling very plausible complete lies. It will be noted that, nevertheless, *C*'s blood pressure modifications were greater, during deception, than those of any other subject. He found lying "easy," but, while deceiving, he felt "like during an exam," "nervous," and "embarrassed." Nevertheless he felt more "tense" during truthful stories, and found them unpleasant, since he felt he could not make the b. p. rise. The apparent great elasticity of this subject's b. p. is to be noted.

Notable Individual Records. No. 2.—*C* introspects that he "worked hard" to raise the b. p. by suppressed laughter, and this is literally true. This "suppressed laughter" involved strenuous muscular contractions all over his body. Such contractions, as would be expected, sent the b. p. up, but it is to be noted that the rise did not follow the form of a lying curve. The b. p. was simply raised 8 mm. and kept there as long as that kind and amount of muscular contraction continued.

No. 7.—More laughter evidently sent the pressure up slightly; but it is probable that its initial level was not the day's norm, and that such laughter had little influence beyond initiating the return to such average level.

C made a mistake in questions 8 and 9, owing to a slip of memory, and recognized that he had made it, but did not correct it lest he be thought to have lied. It is very significant to note that such uncorrected mistake caused no rise of b. p.

Fig. 7

Subject C

T							
2		4		7		12	
	100		110		94		118
End of preparation			←		←	C returns	
	118	10	99	7	102	10 2 2	102 110 108
End of ques. 1–5			←		←	End of preparation	
	118	11	112[1]	10	104	2	106
End of ques. 6–10			←		←	End of recital	
	118	7	110	17	104	2 2 2 2 2 2	116 114 112 110 100 110
End of introspection			←		←	End of cross-exam.	
	96	18	106	15	104	2 2	118 118
						Verdict given	
						2 2	120 118

[1] Suppressed laughter.

Fig. 7—*Continued*

Subject C

L

1		3		5		6		8		9		10		11	
	112		**104**		**118**		**116**		**112**	2 2 2 2	**118 120 122 124 126**	2 2 2 2	**100 104 112** 110 **122**	2 2	130 126 **102**
End of preparation			←		←		←		←		←		←		←
	120		**126**	9	**122**	8	**126**	8	**130**	2 2 2 2 2 2 2 2 2 2	**130 140 146** 136 142 **136 146** 146 140 136	2 2 2 2 2 2	**132 140 144 146** 146 122	2 2 2	**130 138 142**
End of ques. 1–5			←		←		←		←	End of recital			←		←
	128		**128**	10	**130**	16	118[1]	9	130	2 2	120 134	2 2	142 130	2 2 2	**146** 140 144
End of ques. 6–10			←		←		←		←	End of cross-exam.			←		←
	128		**134**	7	124	6	122	12	126	2	138	2	122	2 2 2	130 126 122
End of introspection			←		←		←		←	Verdict given			←		←
	118		102	19	108	14	108	5	118	2 2	122 120	2 2 2 2 2 2 2 2	122 120 118 118 110 112 112 108	2 2	130 132
								Period of quiet.		End of introspection.			←		←
								16	108	2 2	122 116	2	104	2	120

No. 10.—This was highest curve obtained for any subject, and *C* introspected marked *fear*, throughout, which he attributed to the fact that he had not had time to thoroughly construct an alibi.

FIG. 8

Subject D

T					
8		9		12	
	102	2 2 2	92 92 96 94		102
End of preparation			←	D returns	
5	100	2 2 2 2 2 2 2	94 94 90 88 92 94 96	10 2 2	102 90 86
End of ques. 1–5		End of recital.		End of preparation	
7	100	2	90	2	90
End of ques. 6–10		End of cross-exam.		End of recital	
6	90	2	90	2 2 2	94 96 98
End of introspection		Verdict given		End of cross-exam.	
	90	2 2 2 2	96 94 94 94	2	90
		End of introspection		Verdict given	
		2 2	90 92	2	94

FIG. 8 (*continued*)

Subject D.

L																T and L	
1		2		3		5		6		7		10		11		4	
	96		90		84		90		86		90	2 2 2	102 100 88 102	2 2	110 108 100		112
End of preparation			←		←		←		←		←		←		←		←
	102		102		84	7	90	7	102	10	108	2 2 2 2 2	110 128 116 106 96	2	126	6	110
End of ques. 1–5			←		←		←		←		←	End of recital			←	End of ques. 1–5	
	108		104		90	9	100	6	90	7	106	2	96	2 2 2 2	142 106 104 100	8	108
End of ques. 6–10			←		←		←		←		←	End of cross-exam.			←	End of ques. 6–10	
	90		98		94	4	90	8	92	10	106	2	98	2	96	8	122[1]
End of intro-spection			←		←		←		←		←	Verdict given.			←	End of intro-spection.	
	100		98		82	18	88	8	82	13	104	2 2	100 96	2 2	100 100	10	98
														End of intro-spection			
														2 2	96 96		

[1] Lied in answering Ques. 8, 9.

No. 11.—Short but strenuous physical exercise raised the initial b. p. level, but in 4 min. it had returned to 102 or about normal, from which point the "L" curve starts its steady climb.

Subject D.—The stories of this subject were largely negative and as scant as possible. *D* found deception pleasant, and lax, but introspected "*fear* as before an exam," with accompanying "contractions of the diaphragm." He also felt *angry* if forced with questions.

Notable Individual Records. No. 4.—In this record, *D* chose the "T" story, but feared from the first that it would not cover all the questions asked. This feeling seems to have led the subject to lie during questions 8 and 9, and it was only at this point that he felt "diaphragm contraction." It will be noted that at this point the b. p. rose 14 mm. in 8 min., and the experimenter was able to enter a correct judgment, based on this increase, as to the truth and deception of *D*'s story.

No. 5.—*D* tried to "beat the b. p.," by taking no interest in the deception, but it will be noted that the b. p. rose as usual. Also *D* introspected *fear*, and an "alertness despite himself."

No. 12.—Slight physical exertion before *D* came to the experiment at all may have raised the initial b. p. level some what.

Subject E.—*E*'s stories were racy, dramatic, but inaccurate and careless. He was very suspicious of all questions and directions, and had a great desire to outwit the experimenter. While lying he introspected "nervous excitement," inhibitions of ideas due to feeling "like stage *fright*" and "worry" as to b. p., although he found deception very pleasant, and telling the truth uninteresting and indifferent.

Notable Individual Records. No. 4.—*E* added several details to the "T" story, but claimed to regard this just as "telling it in his own words," and introspected "no excitement." From his manner and story, however, he seemed to have a lurking background of vaguely conscious fear that he would be caught up on some of these details, and it will be noted that the b. p. rose very evenly to the slight extent of 6 mm., rather in contrast to its usual more erratic behavior during "T" records.

No. 11.—Initial height of b. p. is due to short, strenuous physical exercise.

No. 12.—Subject while outside after the first b. p. reading sent b. p. up by smoking furiously. He hoped thus to keep it at a high level, and fool the experimenter, but it will be noted that it fell to normal level in 2 min.

FIG. 9

Subject E

T									
1		4		5		7		12	
	108		102		104		108		102
End of preparation			←		←		←		E returns
	98	8	104	4	98	4	108	10 2 2	126 106 92
End of ques. 1–5			←		←		←	End of preparation	
	112	6	108	19	100	9	110	2	106
End of ques. 6–10			←		←		←		End of recital
	96	15	108	7	92	15	108	2 2 2 2	104 102 94 104
End of introspection			←		←		←	End of cross-exam.	
10	100	12	106	7	96	12	108		102
								Verdict delivered	
									94

FIG. 9 (*Continued*)

Subject E

L											
2		3		6		8		10		11	
	99		100		94		94	2 2	90 94 114	2 2	128 114 114
End of preparation			←		←		←		←		←
	100		110	3	98	5	100	2 2 2 2 2 2 2	102 100 98 100 106 92 100	2	114
End of ques. 1–5			←		←		←	End of recital			←
	104		*112*	10	114	6	104	2 2 2 2 2	110[1] 110[1] 102 112 92	2 2 2 2 2	110 132 100 100 126[2]
End of ques. 6–10			←		←		←	End of cross-exam.			←
	106		106	8	102	12	110	2 2	112 104	2	100
End of introspection			←		←		←		Verdict given		←
	100		106	19	94	19	94	2 2	102 88	2 2 2	112 100 112
										End of introspection	
										4	106

[1] Confessed lying. [2] Jury asked betraying question.

Subject F.—Less weight is to be given to this subject's records than to those of any other subject, because of the many elements which would have to be carefully analyzed out.

He told wild, unconvincing stories, and his involuntary movements and stuttering should all have been recorded by a very complicated apparatus if his b. p. records were to be

FIG. 10
Subject F

T						L									
6		8		9		4		5		7		10		11	
	118		114	2 2 2	90 92 94 100		78		96		120	2 2 2	106 108 112 114	2 2	130 118 106
End of preparation			←		←	End of preparation			←		←		←		←
3	110	4	108	2 2 2 2 2 2 2 2 2	102 102 104 102 102 102 102 102 102	7	104	4	102	4	138	2 2 2 2 2 2 2	118 126 116 118 112 112 106	2 2 2 2 2	108 118 110 116 110
End of ques. 1–5			←	End of recital		End of ques. 1–5			←		←	End of recital			←
8	102	12	106	2	100	7	94	9	106	7	128	2	108	2	106
End of ques. 6–10			←	End of cross-exam.		End of ques. 6–10			←		←	End of cross-exam.			←
8	108	12	112	2	100	7	84	12	90	14	112	2	102	2	104
End of introspection			←	Verdict given		End of introspection			←		←	Verdict given			←
10	104	2	92	2 2 2	96 92 94	14	82	12	100	10	110	2	96	2 4	92 100
		Period of quiet		End of introspection										End of introspection	
		10	114	2 2 2	102 96 90									6	100

relied upon as a conclusive test of cerebral factors. During deception he felt "sneaky," cold, "carefree," pleasant, excited and "*frightened*." During "T" stories he felt the "strain"

in trying to be accurate. He inhibited symptoms of extreme nervous *fear* during deception, such as chattering of teeth and trembling of hands.

FIG. 11

Subject G

T				L											
2		4		1		3		6		7		9		11	
	104		90		102		82		98		104	2	92 94	2 2	118 110 112
End of preparation			←		←		←		←		←		←		←
	98	5	86		110		*96*	3	100	3	108	2 2 2 2 2 2 2 2	112 104 100 104 104 112 106 100	2 2	112 110
End of ques. 1–5			←		←		←		←		←	End of recital			←
	84	7	88		100		78	9	118	9	114	2 2 2	104 100 102	2 2 2	118 102 100
End of ques. 6–10			←		←		←		←		←	End of cross-exam.			←
	94	8	90		98		90	7	96	6	118	2	100	2	100
End of introspection			←		←		←		←		←	Verdict given			←
	96	18	86		108		90	13	90	20	106	2 2 2	102 90 90	2 2	112 110
												End of introspection			
												2 2	84 90	2	84

Notable Individual Records. No. 4.—Low general level of b. p. due, apparently, to lack of sleep.

No. 6.—Although the record shows a drop of 16 mm., subject was conscious of having made uncorrected mistakes.

Evidently whatever emotion accompanies this idea does not increase the b. p.

FIG. 11 (*continued*)

Subject G

T and L			
5		8	
	88		102
	←		←
3	88	9	96
End of ques. 1–5			←
10	100	9	102
End of ques. 6–10			←
10	100	10	100
End of introspection		Ques. s. by jury	
20	102	11	112[1]
			←
		9	92

[1] Lied to jury.

No. 9.—Here the b. p. rose to 100 or 102, and was kept there pretty consistently until the subject had entirely finished speaking, not following the form of a curve of deception, but merely exhibiting irregular increases and slight

drops. This peculiar b. p. behavior is to be explained by intermittent body-shaking laughter, and bad stuttering throughout the narration and introspection.

Subject G.—Stories were very plausible, but not very complete and careful. During deception, *G* experienced slight inhibitions, and felt "worried" and "anxious." Telling the truth was harder and less pleasant for *G* than lying.

Notable Individual Records. No. 2.—It will be seen that this shows a typical truth record, and although outside the prescribed conditions of the experiment, the story told was strictly true. After reading the "T" story, *G* thought it improbable. Therefore, to correct an improbable tale he narrated incidents which actually had happened to him, and which, localized at the time and place of the alleged crime, formed a simple and complete alibi. *G* felt throughout that he was correcting a mistake, and telling the real truth, so that this record seems fairly listed as a real truth record.

No. 5.—*G* read the "*T*" story, intending to tell the truth. The second reading shows such was his intent at the moment he finished his preparation. But, yielding to an impulse to improve upon the story given, *G* began to enlarge upon it, with a consequent consciousness of deception and a rise of b. p. At the beginning of the introspection he continued to lie, telling me that he thought he was to use the story given merely as a synopsis but soon laughed and confessed his deception. It will be noted that the b. p. did not fall at the end. In all probability it was much higher at that point toward the end of the introspection where he reached the climax of attempted deception of the experimenter, and had started down toward normal level when the last reading was taken.

No. 8.—This is an excellent example of a simple "T" curve broken by a single lie.

Subject H.—*H* took less actual interest in any sort of work than any other subject and for this reason both the number and quality of his "T" records are significant. His lack of active interest, he introspected, led him to choose the truth 7 times out of 10, and the feeling persisted throughout these "T" records. It will be noted that every one is a downward

curve, varying in drop from 6 mm. to 16 mm. *H* found deception pleasant, but "disquieting," and "irritating."

Notable Individual Records. No. 3.—*H* tried to believe what he said to keep the b. p. down, but his failure is evident in the regular "L" rise of 16 mm.

Fig. 12

Svbject H

T													
2		4		5		6		7		8		11	
	106		96		98		96		106		98	2 2	98 98 98
End of preparation			←		←		←		←		←		←
	104	4	94	3	98	4	92	4	90	6	92	2	88
End of ques. 1–5			←		←		←		←		←	End of recital	
	94	10	92	8	98	9	92	6	92	7	96	2 2	96 90
End of ques. 6–10			←		←		←		←		←	End of cross-exam.	
	92	7	82	9	96	7	90	9	94	14	94	2	98
End of introspection			←		←		←		←		End of ques. by jury	Verdict given	
	96	13	92	10	92	10	96	20	92	3	96	2	92
	Period of quiet										End of introspection		
10	100									14	94		

FIG. 12—*Continued*

Subject H

L								Record Taken During Study	
1		3		9		10			
	92		88	3 2	96 96 96	2 2	108 92 90	3 2 2 2 2 2 2 2 2 2 2 2 2 2 2	102 102 100 106 100 100 100 90 96 96 96 94 96 96 98
	←		←		←		←		
	96		94	2 2 2 2	106 112 108 110	2 2 2 2 2 2 2 2	96 108 108 118 122 120 108 116		
End of ques. 1–5			←	End of recital			←	Total of 30 mins.	
	106		98	2 2 2	106 106 106	2	112		
End of ques. 6–10			←	End of cross-exam.			←		
	[1]		100	2	100	2	120		
End of introspection			←	Verdict given			←		
			104	2	106	2 2 2	112 106 96		
		Period of quiet			←	End of introspection			
		10	86	2 2	104 100	2	92		

[1] Record incomplete.

H's b. p. record taken during study (see also Fig. 1) seems, on the whole, to show more irregularity, and more tendency to hover about two distinct levels than do his "T" records, while it may be said that *H*'s study record was more consistently and evenly downward in trend than those of any other subject.

Fig. 13

Subject I

T						L								T and L	
4		6		9		5		7		8		10		11	
	102		110	2 2 2	126 112 106 112		104		124		114	2 2 2 2	106 106 104 108 106	2 3	110 112 112
End of preparation			←		←		←		←		←		←		←
16	92	8	92	2 2 2 2 2	122 112 110 110 110	6	114	12	128	7	118	2 2 2 2 2 2	112 116 118 110 110 104	2 2 2	118 110 108
End of ques. 1–5			←	End of recital		End of ques. 1–5			←		←	End of recital			←
5	102	7	108	2	116	8	114	8	130	7	116	2	112	2	108
End of ques. 6–10			←	End of cross-exam.		End of ques. 6–10			←		←	End of cross-exam.			←
8	100	7	104	2	118	10	118	12	132	6	120	2	112	2	120
End of introspection			←	Verdict given		End of introspection			←	End of cross-exam. by jury		Verdict given			←
15	104	10	102	2 2 2	116 114 112	13	106	17	120	7	118	2	108	2 5	108 114
				End of introspection						Re-direct exam. ends		End of introspection			←
				2 2 2	122 120 116					2	*124*	5	102	5	108
										Period of quiet					
										17	110				

Subject I.—*I*'s stories were very bare of detail, and told in a prim, precise manner. This was the only subject who preferred truth to deception, and he gave as the reason "that he always found any activity carrying an indifferent feeling tone more pleasant than one stirring up emotion." During

deception he introspected *fear* of coming questions, and "irritation" at difficult queries. He also felt "flustered" and "disgusted" during deception.

Notable Individual Records. No. 11.—Subject opened and read the "L" directions part way through, but changed his mind because he thought of something of his own that he wanted to do. He did not do any physical exercise to raise the b. p., and came back with a "lurking" fear he would be asked if he had read directions. As he thought it over, during the preparation period, this fear became stronger and stronger, he reported, and a corresponding rise in b. p. will be noted. When he was allowed to tell his own story he told the truth without fear, but during the cross examination he was asked if he had opened the envelope, and he replied "No." This was all, and at that point a b. p. rise of 12 mm. in 2 min. is to be noted.

This latter rise is not significant of real deceptive consciousness, but only of one of those isolated emotional associations (of fear) which may occasion abrupt rises of b. p. but do not cause the curve significant of a consistent lying attitude. Such a curve, however, is found in the first part of the record during the preparation period.

Subject J.—*J*'s stories were plausible, consistent, but not ample; and were told in a straightforward manner. *J* gave the ablest introspection of any subject in the experiment. Telling the truth was unpleasant to *J*. He felt "restful," and "could not overcome the feeling of relaxation," but this relaxation was physical and intellectual rather than emotional. Deception was pleasant, exciting, "tense emotionally," with a "feeling of excitement near the heart," frequently "betraying its real affective quality of *fear*."

Notable Individual Records. No. 6.—*J* tried for, and thought he obtained, a "laxness of muscle and attention," but the b. p. shows a good "L" curve, nevertheless.

No. 11.—Initial b. p. was raised by short, hard physical exercise. In this record, *J* introduced an experiment of his own. He chose the truth, but determined to conceal his motive for doing a certain unusual act, and to lie about it if

Fig. 14

Subject J

T		L										T and L	
8		5		6		7		9		10		11	
	114		100		96		100	3 2 2	118 115 116 116	2 2 2	92 92 100 96	2 2	134 122 112
End of preparation			←		←		←		←		←		←
9	108	5	110	4	102	9	112	2 2 2 2 2 2 2	124 130 134 124 132 124 126	1 2 2 2 2 2 2 2	118 96 100 100 100 102 110 102	2 2 2	134 106 122
End of ques. 1–5			←		←		←	End of recital			←		←
12	110	10	116	9	108	7	112	2 2 2 2	120 120 120 120	2 2 2 2	98 106 108 100	2 2	108 116
End of ques. 6–10			←		←		←	End of cross-exam.			←		←
9	110	6	112	8	108	10	118	2	112	2	102	2	118
End of introspection			←		←		←	Verdict given			←		←
2	114	13	102	12	96	18	98	2	102	2 3	110 100	2	120
Period of quiet			←		←		←	End of introspection			←		←
11	114							2 2	108 110	2 2	100 100	2	108

questioned. Throughout the record the b. p. shows a tendency to rise at crucial points, *i. e.*, first, when he told of the act at the beginning of his recital, next where he told of the *motive for another act* at the end of the recital, and finally as he eagerly awaited the verdict of jury and b. p. curve. Nowhere did the falsehood gain sufficient importance in the

story to bring about a full deceptive attitude, but it constantly tended to do so whenever the fear in the background of consciousness was touched by associations, or crept toward the focus in expression. In short, this curve represents a story told in fear that the witness will be obliged to lie, with a final fear that the single point of deception may have been detected after all.

D. *Interpretation of Individual Results*

Benussi, as above mentioned, made scarcely any attempt to analyze or to psychologically describe the deceptive consciousness. He does report, however, that his subjects found the work of lying hard, disquieting, and unpleasant; and that they introspected "tension of attention, excitement, and discomfort." As reported, this introspection does not seem illuminating, nor does it agree with the introspection of the subjects in this experiment. It will be noted in the individual results reported above, that all subjects, with the exception of Subject *I*, reported the lying more pleasant than telling the truth. Moreover the pleasantness of the whole attitude, or consciousness, seems to depend upon the added interest of the whole proceeding, as an adventure is more pleasant than routine, and seems also to depend upon the success or failure of the attempt to deceive at any particular point. In this, the deceptive consciousness seems to resemble every other complex state of mind, and does not admit pleasant or unpleasant affective tone as a crucial criterion, or even as a consistent constituent. Nor does "tenseness" serve as any better indicator of deception. Seven subjects introspected tension, and four further designated this feeling as "affective tension." These four were asked, at each experiment, to record the height of this feeling and in no single instance did the "tension" climax have any correlation whatsoever with the climax of the b. p. curve. The tension element really consists of what Benussi calls "tension of attention," or, as it would ordinarily be called, concentration of attention on the task before the subject. It has been made clear by experiment in other fields that concentration of attention involves a

certain involuntary setting of the muscles, and very probably general contractions of large groups of these muscles. Yet these same contractions, due to concentration of attention, occur during study, or during the other forms of mental work used in the preliminary intellectual work records, and, in fact, exactly the same sort of concentration with a feeling of "tension" was introspected during several of the "T" records. It is significant that in these "T" records the b. p. did not fall, but remained on an almost exactly even level throughout. Thus we must recognize a certain tonicity of involuntary muscles due to concentrated attention as an almost constant, if not invariable, concomitant of deception; but we must also recognize that the utmost function of such tonicity is to keep the b. p. from falling, and that concentrated attention is by no means peculiar to deception. We must then seek further for the essentially characteristic constituents of deceptive consciousness.

Fear and *anger*, as mentioned above, are the only two emotions which could produce *moderate* increases of b. p. and since the records above show just such moderate but persistent b. p. changes, it would seem probable that one or both emotions constitute the true key to the mental state during lying. It will be noted, from the individual introspection given above, that every subject introspected some complex emotional state containing the element of fear, while many designated the feeling as simple fear. "Feeling of responsibility," "fear of awkward questions," "nervousness as in an exam," "worry," "sneaky feeling," and feeling "flustered" all point inevitably to fear as the common denominator and chief factor of all introspections during deception.

Is *fear*, then, the sole emotional element in the deceptive consciousness? It seems probable that, during a majority of the deception, it is.

Five subjects, however, introspect "irritation" and *anger* at certain points in their false testimony. Outward signs of anger appeared, in all subjects, whenever they were outwitted into betraying themselves under cross-examination, when they gave their case away by careless inconsistencies, and

occasionally when they considered that they were pressed too closely with questions. From the introspection of the subjects in this experiment, then, it plainly appears that fear always, and anger when in immediate danger of detection, are the characteristic emotional factors betraying deception through the b. p.

What, then, is the psychological organization and mode of operation of fear and anger during deception, and what is their relation to the other conscious elements then present? The stimulus to fear is, of course, a central situation mirroring a relation of *danger* between subject and environment. Professor Cannon substantially proved that, with regard to the primary factor of fear,[1] "the natural response is a pattern reaction, like inborn reflexes of low order, such as sneezing, in which impulses flash through peculiarly coöperating neurone groups of the central system, suddenly, unexpectedly, and in a manner not exactly reproducible by volition."[2] That this central response of fear may occur instantly in reaction to sudden danger is a matter of everyday experience, and that ensuing b. p. changes are scarcely less instantaneous is evidenced by records like No. 11 of Subject *I* where, at the telling of a single monosyllabic lie, the b. p. rose 12 mm. in 2 min. If, then, the b. p. response so immediately follows the creation of a dangerous situation, we would expect to find either an initial increase as great as the b. p. elasticity of the individual permitted, with b. p. remaining at this level throughout the deception; or simply a series of sharp isolated rises at each new lie. In a few of the above results we find "L" records exhibiting the last-mentioned tendency, but far more frequently we find the curve designated above as the significant lying curve. Another factor, therefore, must be at work. The introspection indicates that this factor is to be found in the attempted voluntary control of the fear impulse. The attempt took a distinctly individual form in every case, but the common factor in all the methods employed as reported by introspection was the attempted elimi-

[1] And also of pain and anger.

[2] W. B. Cannon, "Interrelations of Emotions," *Am. Jour. Psy.*, XXV., p. 281.

nation of fear from consciousness by a fixation of the attention on the purely intellectual elements of the story. Thus a significant lying curve is a function of the struggle between the involuntary impulse to express fear in response to awareness of danger, and the voluntary focusing of attention to exclude the fear from consciousness. As the ideational elements of the deception become more and more complex, the awareness of danger becomes more and more firmly established in the foreground of consciousness, and, as the stimulus is thus enhanced, the "natural response" of fear becomes stronger and stronger. In some cases the fear impulse probably never entirely breaks away from the restraint imposed by voluntary inhibitions, but in other records we see evidence that, at the *danger* climax, the fear impulse is wholely uncontrolled. The close correlation between the height of the "L" curve and the climax of the intellectual task of deception[1] evidentially substantiates the nature of this danger climax, and points to exactly that gradual return to normal b. p. level which actually appears in the records. This gradual decrease of fear symptoms is due not so much to strengthening of voluntary control, as to decrease in the force of the fear stimulus, *i. e.*, awareness of danger in the case of most subjects. The voluntary control seems to decrease with the necessity for such control and for this reason questions put at the very end of the cross-examination, or by the jury, may cause the fear impulse to run its momentary course unimpeded, betraying itself in short, pronounced b. p. increase. In the same way we may now explain the b. p. behavior when a subject comes to the experiment with the knowledge that he may be obliged to lie if certain crucial questions are asked concerning "guilty" acts. After the first question is safely past, the telling of the truthful story steadily removes the central stimulus to fear, and, correspondingly, the voluntary control is allowed to relax, with the result that if, later, the awareness of danger actually rises to the focus of attention, the resultant fear impulse is unchecked by volition (see record No. 11, subject *J*).

[1] See page 18 above.

Exceptional subjects may, however, retain in consciousness their voluntary inhibitions to fear impulses despite the cessation of those impulses, and in such cases low, *significant* curves of deception take the place of short, sharp, isolated rises.

"Fear is a reaction aroused by the same objects that arouse ferocity," says James ". . . the question which of the two impulses we shall follow is usually decided by some one of those *collateral circumstances* of the particular case."[1]

I would suggest that the collateral circumstance which always turns fear into rage is the occupation of the focus of attention by the awareness that there is no escape from the danger impending. This may occur, during deception, either by some sudden betrayal, or by the victory of fear in winning its way to conscious motor expression and so betraying the lie. In either case, the anger impulse supersedes the central fear reflex, and rage is registered in consciousness. But, since both emotions are expressed through the sympathetic nervous system, the visceral changes which have been taking place during fear continue during rage, and the b. p. level will merely depend upon the strength of the anger impulse aroused. If (as with Subject A, Record No. 5) the anger is slight, and of short duration, it will almost immediately allow the b. p. to drop as during a "T" record; if, however (as with Subject *E*, Record No. 10, No. 11), the anger felt is, at least momentarily, stronger than the fear previously in consciousness, a rise in b. p. will take place. In a situation such as that produced by deception in real life, it is very probable that rage at detection or at self-betrayal would be far more intense and of much longer duration than under the artificial conditions of the laboratory. Finally, there is little possibility that, in this experiment, emotions other than rage and fear were aroused to a sufficient degree of intensity during deception to break over into the sympathetic nerve channels and so affect the b. p., and probably in actual experience such contingencies would be of almost equally rare occurrence.

It may be noted at this point that, if we resolve the crucial

[1] "Principles of Psychology," II, p. 415.

deceptive factors into fear and rage, we have a seeming anomaly in the almost unanimous introspective report that deception was *pleasant*. But it was the *whole experience* of the deception that the subjects found pleasant, not the isolated fear, and perhaps not the actual moments when the fear was recognized as such. Men prefer bluffing at poker to playing a conservative hand, and the explanation would seem to be that excitement is more pleasant than quietude, any emotional experience being preferred, perhaps, to a purely intellectual activity. Where, as in actual court work, life and happiness might hang on the success of a deception, it is much more doubtful whether the whole situation during deception would be more pleasant than while telling the truth. Probably, in such cases, both would be disagreeable.

Whether the experience of fear *can* be pleasant per se is a more difficult question, which need not be decided here. The popularity of such amusements as "scenic railways," the sole attraction of which lies in the *fright* on the steep inclines, suggests that fear may be pleasant—at least retrospectively.

Of the "T" records, 76 per cent. show downward b. p. curves, 62 per cent. showing final return of b. p. toward its initial level. On the whole this unlooked-for result seems too persistent to be entirely disregarded. Certainly no significant emotional element appears in the "T" introspection, and the only common factor there seems to be a very general mental and physical feeling of passivity. Thus the effort necessary to maintain attention on the task in hand is prominent in consciousness, and the subject remains indifferent to mistakes, or failures of memory. A general passive mental attitude of this sort might affect the b. p. in two ways: (1) By general relaxation of all muscles, and consequent lack of cardio-accelerator impulses, with the possibility of actual inhibitions at the accelerator center; (2) by vaso-dilations throughout the body such as occur during sleep,[1] with consequent lowering of diastolic b. p. and indirect reduction in systolic pressure. Such influences would, in short, tend to remove several of the ordinarily acting physiological factors

[1] See Howell's "Physiology," 5th ed., pp. 255ff.

which determine the subject's average b. p.; and would not prevail against any positive influences which might be suddenly introduced. That the uniformity of "T" curves is due to some such general negative influence is indicated by (1) the fact that it does not take effect in 24 per cent. of the "T" records; (2) the fact that it is clearly overcome in a few records by positive influences such as muscular contractions; (3) the fact that the average drop in b. p. during "T" records was only 8 mm., as contrasted to an average increase of 16 mm. for all "L" records, and (4) the fact that the "T" curves are much more irregular than the "L" curves, and subject to inexplicable variations as large as that record's minimum decrease. Thus it appears fairly certain that the downward "T" curve is a function of one or more general, negative influences, but it is much more difficult (and perhaps less profitable) to discover which of the great possible numbers of such influences determine the result recorded in the "T" records.

E. Judgments of Experimenter and Jurymen

A total of 107 records were taken in this experiment. 10 of these records were "T and L" records, and 3 of the "T" records were influenced by muscular contractions. The experimenter, basing his judgment entirely upon the b. p. behavior, made 103 correct judgments and 4 erroneous ones. Of the latter, 3 mistakes were made on "T and L" records, and 1 mistake on a "T" record influenced by muscular contractions. The b. p. judgments were, then, 96 per cent. correct.

Benussi reported that he found two distinct classes of liars divided upon the basis of their success in deceiving his "trained observers." He further reported that the breathing of the successful liars was more modified during deception than was that of those unsuccessful at deception. No such result was obtained in this experiment, however, as the following table will show. Only the judgments of those men who sat regularly in the juries and whose judgments were therefore susceptible of comparative study, are recorded in Fig. 15.

FIG. 15

Jurors' Judgments

Subjects...	*A*		*B*		*C*		*D*		*E*		*F*		*G*		*H*		*I*		*J*		% R.
Jurymen...	R.	W.	R.	W.	R.	W.	R.	W.	R.	W.	R.	W.	R.	W.	R.	W.	R.	W.	R.	W.	
1												1							2	4	.28
2																			4	3	.57
3					5	7															.41
4					7	3			5	3							5	2			.68
5													6	4							.60
6		1											2	7							.20
7							3	7											1	3	.28
8							4	1													.80
9							2	1			5	1							3	1	.76
10	2	1											2		6	4					.66
11	.		1	1							3	4			1	2	1	2			.40
12											4	4									.50
13	1	3							2	2							2	4			.35
14																	2	5			.28
15		4	1	5																	.10
16			4	1													3	2			.70
17			4	3						2											.44
18	4	1			2																.85
19	2								5	1	1	4									.61
20									2	6		3									.18
% W.	.53		.50		.42		.50		.50		.57		.52		.46		.54		.52		

Note.—"R." = Right and W. = Wrong, judgments. Per cent. R. = Per cent. of juror's correct judgments. Per cent. W = Per cent. of subject's successful deceptions of the jurors.

It is clearly evident, from the above table, that the jurors, not the subjects, are the ones to be divided into successful and unsuccessful classes. 7 of the subjects were from 50 per cent. to 57 per cent. successful in deceiving the jury, but these percentages are obviously too low to be significant. Jurors 4, 9 and 10, on the other hand were clearly successful at "sizing up" the subject, all three being consistently successful in judgments made upon three different subjects. Other jurors, such as 1, 6 and 20, are clearly poor judges of men, and, although the classes of successful and unsuccessful jurors approach within 20 per cent. of each other, only No. 12 is squarely on that neutral line of 50 per cent. about which the percentages of the subjects so consistently hover.

F. Series D, Checking Series.

This series followed exactly the method of Series *B*, an extra story being prepared and used throughout the series. A

juror was selected without any warning, and requested to take the subject's place for that day, while the subject took his place in the jury. The chief importance of this series lies in testing the practical value of the b. p. record as an indicator of deception. These subjects could have had no fear or other emotion associated with previous experiences as subjects in that experiment, and were in exactly the position of any naïve witnesses who might be called upon anywhere to testify. It will be noted that 5 clearly *significant* "L" records were obtained, that the "T" records were all of a persistent downward tendency and that, when juror 9 lied a little in answer to the last question, he was immediately betrayed by the b. p. All b. p. judgments in this series were correct.

Fig. 16

Extra Story, Series D

T						L										T and L	
1		2		3		4		5		6		7		8		9	
	116		108		118		126		110		130		118		124		112
End of preparation			←		←		←		←		←		←		←		←
4	114	6	98	4	110	3	140	10	122	8	142	10	134	10	138	6	108
End of ques. 1–5			←		←		←		←		←		←		←		←
7	112	7	102	6	108	8	144	7	126	6	128	6	126	8	140	7	106
End of ques. 5–10			←		←		←		←		←		←		←		←
12	116	11	106	9	116	8	138	6	118	8 2	130 122	10	116	9	128	9	118[1]
End of cross-exam. by experimenter			←		←		←		←		←	End of cross-exam.			←	End of cross-exam. by experimenter	
2	116	1	102	3	120	2	140	3	120	2 6	132 128	15	108	11	110	1	112
End of cross-exam. by jury			←		←		←		←		←						←
6	114	12	100	12	118	12	124	3	120	6	118					14	102

[1] Lied at Ques. 10.

With reference to the practical application of the b. p. test, a final individual experiment is of interest. Mr. Dewey, who had witnessed several of the experiments, proposed to tell two stories of his actions on a certain afternoon, both stories to be objectively true, but one set of actions to be those which he performed during another afternoon than the one selected. The curves in Fig. 17 show the result. The "L" curve, while not great in height, is perfect in form as a significant deception curve, and contrasts sharply with the truth curve which,

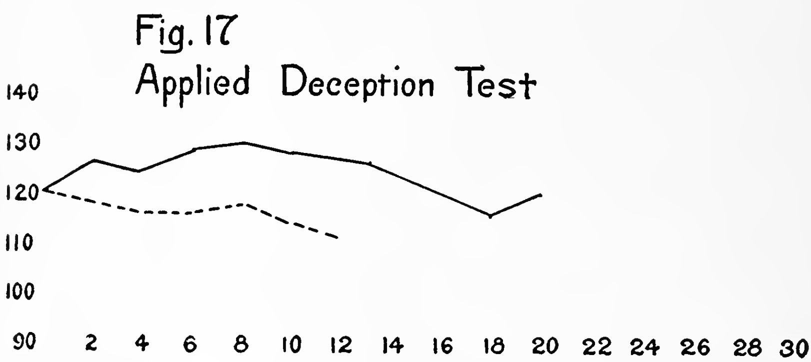


FIG. 17.

starting at identically the same b. p. level, drops slowly and evenly 8 mm. The experimenter passed over, in writing, the correct judgment at the end of the fourth b. p. reading of the second curve.

5. CONCLUSIONS

1. The behavior of the b. p. does not act as the least indicator of the objective validity of the story told by any witness, but it constitutes a practically infallible test of the *consciousness of an attitude of deception.* Mere awareness of a mistake, even if the mistake is uncorrected, or the mere addition of trifling details, even if the subject is conscious of such additions, will not constitute that mental situation which is the necessary stimulus to fear, and will not, therefore, cause the b. p. to rise.

2. The *significant* curve of deception differentiates a story the foundation of which is false from a story mostly

true, but containing one or two substantial lies. The sudden sharp, short rises of b. p. betray these substantial lies in an otherwise true story. It seems probable that, if a truthful witness became violently angry at some chance question of the examiner, or if he suddenly saw his worst enemy glaring at him, gun in hand, in the court-room, his b. p. would suffer a short, abrupt rise, but if such extreme outside influences are avoided, all major b. p. modifications would seem to depend upon the deception elements of the story itself.

3. The b. p. record during testimony might be made practical use of as an indicator of deception if the test embodied the following features:

(*a*) Two records must be taken as in the test on Mr. Dewey, the story told during one record being truth within the knowledge of the examiner.

(*b*) The examination should be private, with carefully controlled conditions, and means at hand for recording involuntary movements, muscular contractions, and sudden or suppressed laughter.

(*c*) The record should be interpreted by a psychologist experienced in this particular line, and should be scrutinized with careful reference to the construction and subject matter of the story, the record of the manner and muscular contractions of the witness, and above all it should be compared minutely with the record known to be symptomatic of that individual's consciousness while telling the truth.[1]

[1] The writer expresses his gratitude to Professor Münsterberg and Professor Herbert S. Langfeld, to his assistants, E. H. Marston and T. Ramsdell, and to the Harvard men who served as jurors.

Journal of
Experimental Psychology

Vol. II, No. 3. June, 1917

THE EFFECT OF RAPID CHANGES OF WORK ON THE RATE OF PERFORMANCE

BY J. CROSBY CHAPMAN

Western Reserve University

Several papers[1] have been published recently investigating the problem of the amount of work produced in successive periods of short duration. The results prove conclusively that in several of the simple mental functions such as addition, opposites naming, color naming, cancellation, etc., there is distinct evidence of the presence of a so-called "initial spurt." That is during the first two or three half-minute periods, the work produced is sometimes as much as 20 per cent. in excess of that performed during the later periods. This factor of initial spurt is so great that the variation in the amount of work produced in successive periods of short duration seems to be worthy of careful quantitative study. In the present paper a cognate problem is attacked, namely the effect of rapid changes of set on the rate of performance of work.

Whenever the individual turns from one type of mental operation to another, a certain change of "set" or "attitude" is required. In complex operations the number of changes of set necessitated must be very great. Taking two mental operations such as those required in a simple addition and an ordinary cancellation test it is desirable to measure the extent to which speed of work is affected by rapid changes of set. This can be accomplished by comparing the rates of work under the two following conditions: (1) when only one oper-

[1] Poffenberger and Tallman, *Psych. Rev.*, Sep., 1915, p. 371; Chapman, *Jour. Educ. Psych.*, Sep., 1915, p. 419; Chapman and Nolan, *Amer. Jour. Psych.*, April, 1916, p. 256; Thorndike, *Amer. Jour. Psych.*, Oct., 1916, p. 550.

ation is performed for a certain period; (2) when two operations are performed alternately for the same period. It was to make the comparison that this study was undertaken.

The general arrangement of the experiment was such as to test the efficiency of a group of subjects on five occasions. On each of these occasions the work was divided up into six trials, each of three minutes; during each of these three-minute trials the work at one operation or both, as the case might be, was continuous. In the table below the arrangement of the experiment during each of the six three-minute trials is shown; *A* and *C* signify addition and cancellation respectively.

Trial	Half-minute Periods					
	1	2	3	4	5	6
I	*A*	*A*	*A*	*A*	*A*	*A*
II	*C*	*C*	*C*	*C*	*C*	*C*
III	*A*	*C*	*A*	*C*	*A*	*C*
IV	*C*	*A*	*C*	*A*	*C*	*A*
V	*C*	*C*	*C*	*C*	*C*	*C*
VI	*A*	*A*	*A*	*A*	*A*	*A*

Thus during trial 1, addition was performed steadily for three minutes; during trial 2 cancellation; while during trial 3 addition and cancellation were done alternately. The only difference between trials 3 and 4 and the other trials was that in the two former there were two types of work, involving a change of set every 30 seconds, while in the rest of the trials only one operation was involved throughout.

This alteration of work was accomplished by having addition on one side of the test blank and cancellation on the other. At a signal every half minute the subjects reversed the sheets and passed from the one operation to the other. The point reached in the case of addition was marked by putting down the partial result of the column not completed. The Thorndike addition and the Woodworth and Wells' cancellation sheets were employed. In order to eliminate every factor except the change of set, when the subjects did the continuous cancellation and addition, they were made to reverse the sheets every half minute and proceed on the reversed side with the same operation. It may be said that

the reversing of sheets was practised during the preliminary periods, so that it became almost automatic.

The subjects were pupils of the junior class of Glenville High School, Cleveland. The nature and operation of the tests were carefully explained to the pupils, but the object of the experiment was not revealed. Particular care was taken, during the conduct of the test, to see that the sheets were reversed at the time when the signal was given; it was pointed out that any delay, accidental or purposive, might vitiate the experiment.

The basis of scoring was ten for each column added correctly, a score on the same basis being given for part of a column added. The score in the cancellation test was on the usual arbitrary basis: 2 (number cancelled correctly)—2 (number omitted)—3 (number wrongly marked). In the following tables the scores are recorded for the total number of individuals taking the test on the particular day.

TABLE I

Day	Period I						Period II						Period III					
	A	*A*	*A*	*A*	*A*	*A*	*C*	*C*	*C*	*C*	*C*	*C*	*A*	*C*	*A*	*C*	*A*	*C*
1 (18 indivs.)	390	367	367	339	372	322	760	724	652	586	684	660	497	736	405	740	447	752
2 (29 indivs.)	755	667	637	596	673	626	1,444	1,450	1,378	1,458	1,312	1,384	728	1,530	657	1,368	672	1,366
3 (24 indivs.)	459	464	526	531	492	550	1,282	1,290	1,168	1,160	1,218	1,228	647	1,284	503	1,220	567	1,246
4 (24 indivs.)	727	605	552	546	546	542	1,582	1,448	1,296	1,240	1,354	1,250	705	1,424	602	1,278	576	1,340
5 (25 indivs.)	580	598	660	576	596	549	1,554	1,376	1,378	1,360	1,412	1,428	586	1,490	672	1,404	666	1,482

Day	Period IV						Period V						Period VI					
	C	*A*	*C*	*A*	*C*	*A*	*C*	*C*	*C*	*C*	*C*	*C*	*A*	*A*	*A*	*A*	*A*	*A*
1 (18 indivs.)	792	456	746	443	816	469	926	898	806	792	760	778	398	367	330	346	346	343
2 (29 indivs.)	1,580	660	1,404	679	1,502	726	1,628	1,572	1,478	1,472	1,462	1,446	720	626	575	573	609	603
3 (24 indivs.)	1,440	576	1,128	583	1,312	597	1,394	1,352	1,232	1,178	1,294	1,246	553	529	531	474	506	504
4 (24 indivs.)	1,472	647	1,370	624	1,262	655	1,482	1,398	1,318	1,290	1,224	1,278	645	556	563	532	526	515
5 (25 indivs.)	1,536	719	1,402	662	1,418	633	1,664	1,508	1,450	1,370	1,438	1,558	674	619	576	560	608	593

Using these data it is possible to compare the amount of addition performed in the periods when there is a change of set with the amount produced when there is continuous work at the same operation. The same results can be deduced in the cancellation test. It will be noticed that in the treatment of the data, owing to the arrangement of the periods, the practice factor is eliminated. The scores in Table II are

worked out on the percentage basis; 100 being the scores obtained during the periods when there is continuous work in the same operation.

TABLE II

		Average	P. E. of Average
Addition	Continuous work	100	1.3
	Change of set	114	1.5
Cancellation	Continuous work	100	1.6
	Change of set	100.6	1.8

From this table it will be seen that the subjects were 14 per cent. more efficient in adding when they were alternately performing cancellation, than when they were adding continuously. On the percentage basis P.E. of difference = 1.3, that is, the difference is nearly eleven times the unreliability. In the case of cancellation the effect of change of set is negligible, for the same amount of work is produced when the operation is continuous as is obtained when there is change from this operation to that of addition, every half minute.

We are thus faced with having to explain the fact that while a change from addition to cancellation, and back to addition, etc., does not seem to exert any effect on the rate of cancellation when compared with continuous work, yet in a mental operation such as addition it causes an actual increase in efficiency of 14 per cent. What is the difference in the two functions which causes the effects of a change of set to be so different? The factor of increased interest could not account for this differential result. The reason is probably similar to that suggested for the explanation of the initial spurt effect, namely that it is a matter of interference. In the case of continuous cancellation, chiefly a motor test, the previous cancellations do not have to be given time to fade away to any marked degree; little has to be banished from the mind when the subject proceeds to the next cancellation, whereas in addition, chiefly a mental function, when the passage is made from one step to another all previous results, except the one necessary for the immediate calculation, must be forgotten. Thus if we imagine an individual commencing with no figures in the mind at all, it would seem reasonable that the individual would go faster in the first few seconds, than in the later seconds, when the mind has of necessity to be given time for

the previous totals to fade away and obliterate themselves. When the subject passes from cancellation to addition, the subject as far as addition is concerned is better off than when passing from addition to addition; such an explanation is necessary to account for the results obtained.

This hypothesis is strengthened by the evidence obtained in the following experiments. Here the conditions were changed so that instead of having cancellation alternating with addition, a period of complete rest was substituted. The distribution of time is shown below.

Trial	Half-minute Periods								
	1	2	3	4	5	6	7	8	9
I.	*A*	*R*	*A*	*R*	*A*	*R*	*A*	*R*	*A*
II.	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*
III.	*A*	*R*	*A*	*R*	*A*	*R*	*A*	*R*	*A*

A = *addition*. *R* = *rest*.

The procedure of the experiment was similar to that employed in the previous experiments, except that on the reversal of the sheet during trial 1 and trial 3, no adding was done. The subjects were 24 students of an undergraduate class. In this experiment the average amount of work done during trials 1 and 3 (periods 3, 5, 7, 9) is compared with the amount done during trial 2 (periods 3, 5, 7, 9). These results are shown in Table III.

Table III

Trial	Half-minute Periods			
	3	5	7	9
I.	31.3	28.3	32.1	33.1
II.	25.8	26.8	26.3	26.3
III.	35.1	31.5	31.1	30.1

The comparison of the amount of work produced under (1) conditions of change every half minute from rest to addition; (2) continuous work; is shown in the next table reduced to the basis of 100 for the continuous work.

Table IV

		Average	P. E. of Average
Addition	Continuous	100	.3
	Broken by half-minutes rest periods	121	1.6
	Unreliability of difference	1.6	

Here it will be seen that the complete rest alternating with addition, has a distinctly greater effect on the amount of addition performed than the change from cancellation to addition affords. This would agree with the explanation that the decrease in product produced after the first half minute is caused by an interference produced by previous work, for in this case complete rest should have a greater effect than the mere change which is afforded by passing from cancellation to addition.

The results of this paper give strong support to a previous hypothesis,[1] that the phenomenon of initial spurt is probably due to an interference effect which results from continued work. Thorndike[2] also gives this as an explanation of his own results which indicate a higher efficiency at the commencement of the work than during its steady progress. If these contentions are true, we shall have to avoid the phrase initial spurt, which seems to imply, not merely a power of producing a greater external product, but also a greater internal initial effort, and replace it by a term which contains the idea of initial maximum performance produced by lack of interference.

To summarize:—This experiment indicates that rapid changes of set, every half minute, from cancellation to addition, result in a greater combined speed of work than is found when either of the operations occur separately. Far from the change of set involving, as might be expected, a certain time lost in adjustment, it effects an increase of speed. The explanation is probably found in interference. This suggests that the last named factor occupies a more important place than is supposed in mental operations, and that its effects should be measured over much shorter periods of time.

I wish to express my indebtedness to Miss Mansfield who conducted and scored a portion of the experiment, and to the authorities of the Glenville High School for the facilities which were afforded.

[1] Chapman and Nolan, *Amer. Jour. Psych.*, April, 1916, p. 260.

[2] Thorndike, *Amer. Jour. Psych.*, Oct., 1916, p. 565.

APPARATUS FOR RECORDING CONTINUOUS DISCRIMINATION REACTIONS

BY H. C. McCOMAS

Princeton University

A series of discrimination reactions has been used in connection with a number of investigations having widely divergent interests. The records obtainable in time and accuracy for such reactions afford side lights upon a variety of processes. It is not surprising, therefore, to find some type of discrimination apparatus, from Seashore's "Psychergograph" of fifteen years ago to the numerous forms of today, employed in experiments on mental efficiency, concentration of attention, influence of distractions, motor impulses, etc.

The devices used are adapted, of course, to the immediate purposes of each experimenter, though the general principle is the same; namely, to present several different stimuli and to record the several possible reactions continuously. To obtain reliable results in this work a number of conditions must be squarely met. That part of the apparatus with which the subject has to do must be noiseless, with no moving parts that afford distractions. The stimuli must be constant in intensity and sufficiently varied in the order of presentation to obviate anticipation. The type of response must admit sufficient variation to preclude the quick formation of habit. Fatigue in sense-organ and reacting muscles must be reduced to a minimum. Time records should be obtainable for both right and wrong reactions during variable periods.

In addition to these indispensable requirements it is desirable to have an apparatus to which all subjects become quickly adapted. This shortens the practice curves. If the reaction is simple and the discrimination is not complicated, the training period is reduced; and the results of different subjects are more comparable. For a large class of subjects it is a decided advantage to have work which is interesting. The kymograph

records are long and should be easily legible. This can be partially accomplished by markers with wide amplitude and set close together. As this type of apparatus is not manufactured it is desirable that it should be easily and cheaply made.

The apparatus to be described meets these requirements more fully than the usual types. It is capable of many adaptations. The work in which it has been most used was in obtaining ten minute records of continuous discriminations. The procedure was briefly this: the subject sat in one room, with the exposure screen and reacting keys; the experimenter and the rest of the apparatus were in another room. The subject reacted to four differently colored lights, by pressing different keys. He learned which keys extinguished which lights before the records began, and then he reacted as quickly as he could with the fewest possible false reactions. The instant one colored light was extinguished another appeared; the order of their appearance was miscellaneous. The subject therefore controlled the rate of the exposures. The records were read for ten second intervals and showed that the average subject makes from eight to thirteen correct reactions in such an interval and from no false reactions to three or four. The time of reaction and the number of false reactions increase as the test goes on. The records show typical differences in subjects and characteristic forms of reaction for some subjects at different times of the day.

In describing the apparatus it may be conveniently considered in three parts: first, the stimulus and reaction table; second, the automatic switchboard; third, the recording instruments.

Upon a small table is placed a green cardboard screen with a small ground glass window in its center. In a rack behind the window are four Mazda, six-volt bulbs. In front of these are four different colored gelatine strips. The bulbs are connected with a Hyray battery. This service brings the filaments to incandescence in about .05 sec. after a circuit is closed. Four telegraph sending keys are placed in front of the screen upon the table, and in such a position that they

are easily seen in indirect vision when the eyes are fixated upon the window (Plate I.). They are wired to make two circuits; one by the contacts at the far end of the rocking arm and the other by the contacts at the near end (Plate IV., *LC* and *MC*). The reaction movement breaks the circuit in which the light is made and at the same time it makes a circuit which actuates a marker on the kymograph. Thus

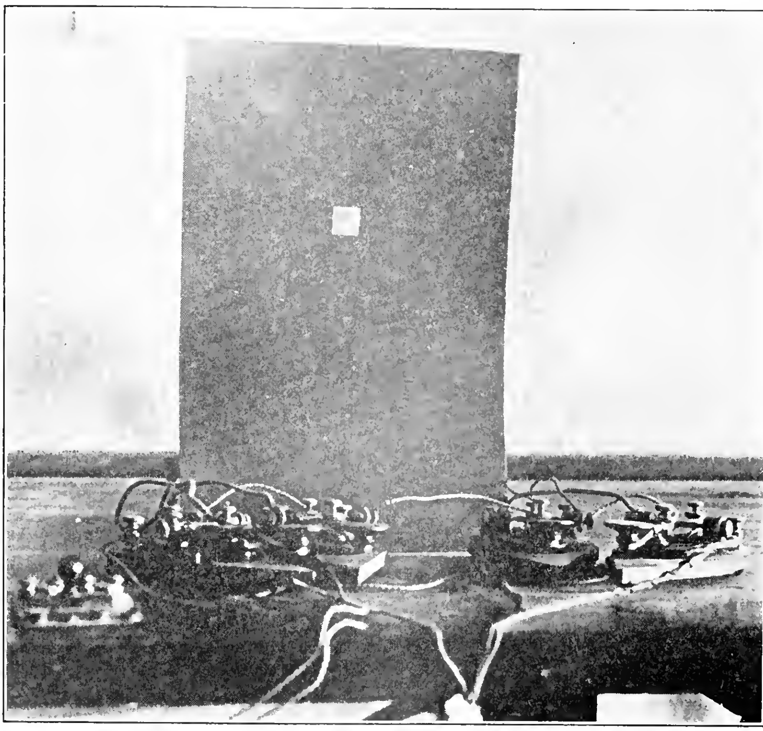


PLATE I.

the light is extinguished and a reaction recorded by one movement.

An automatic switchboard is used to remake the light circuits. This device is easily and cheaply made by fastening a 'secondary' electric clock opposite a board on which there is a circle of radiating copper contacts (Plate II.). The 'secondary' is a simple mechanism by which the minute

hand of the clock is moved forward with each break of the current running into the magnet. By placing a trailer contact on the minute hand and adjusting it to pass over the radiating copper contacts, it is possible to have the hand move forward to a new contact and make a new circuit whenever a circuit is broken by a reaction. The sixty copper contacts are fastened upon the supporting board as radii coincident with the minute divisions of a clock face. From the contacts wires pass through the board and rest in four bowls of mercury, fifteen wires to the bowl. From each bowl a wire runs to

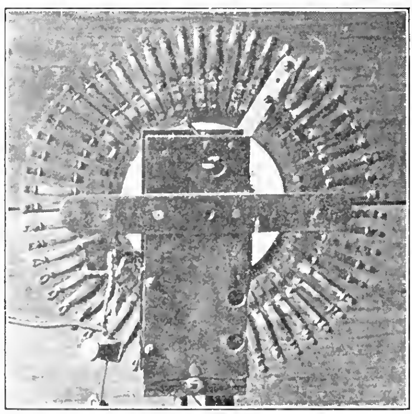


Plate II.

each of the four keys. When a light circuit is broken at one of the reaction keys, such as *LC* in key K_4, Plate IV., the light L_4 is extinguished, the current ceases to pass through the coils in magnet *M*, this releases armature *A* and the arm *AR* swings counter clockwise moving the hand *H* forward to another contact. The time required to pass from one contact to the next is .08 sec. This can be shortened by placing the copper strips closer together, and by adding to the weight at *N*.

The records are taken upon a kymograph. A Jaquet marker gives the time in seconds. A marker in circuit with the armature of the clock registers each correct reaction and a marker in circuit with the reacting keys scores all the reactions

made. These markers are constructed from two telegraph sounders with long pliable copper strips to give a wide amplitude (Plate III.). They are set close together, the all-reactions registering immediately below the correct reactions.

The wiring of these three divisions of the apparatus can be easily understood by reference to Plate IV. There are two independent battery circuits, the light circuit and the marker circuits. The light circuit may be traced from the battery B_1 to the four bulbs $L_1L_2L_3L_4$ and from these to the four reacting keys $K_1K_2K_3K_4$. If all the circuits stood open except K_4, the current would pass through LC into the wire R_4 running to the automatic switchboard. There it would enter

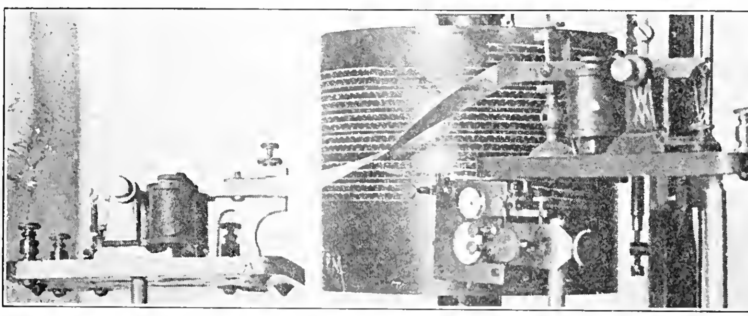


Plate III.

one of the four mercury bowls from which it would pass to the radiating copper strips CS. On one of these the trailer contact H is resting. The current, therefore, passes through H down to the magnet M, actuating armature A and thence back to the battery.

The other two circuits lead from the battery B_2, they are the marker circuits. With the reaction to each light the current in M is broken and the lever AR moves about the axle P. With each movement a contact is made at N and the marker Y registers on the kymograph KY giving a record of the right reactions. The marker Z is actuated by closing such a contact as MC in K_4. The wires F_1 and F_2 connect the keys with the battery and the marker.

If it is desired to give a longer series of discrimination

reactions than sixty at a sitting it may be advisable to put a sliding switch in $R_1R_2R_3R_4$ and thus shift the order in which the lights appear in $L_1L_2L_3L_4$. However, many subjects do

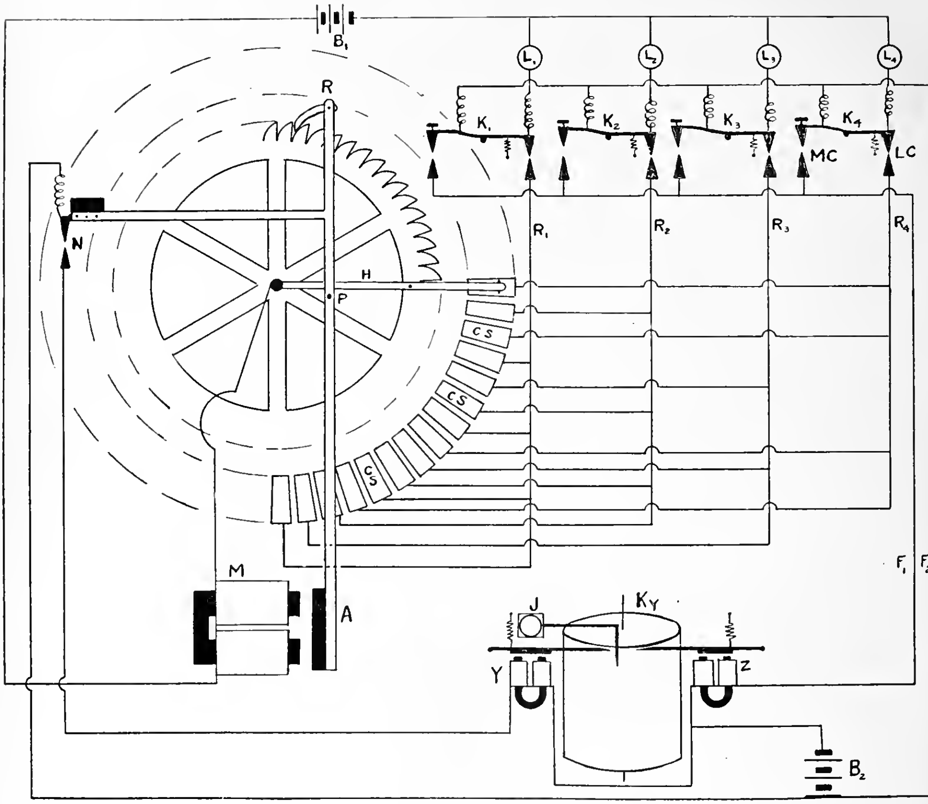


PLATE IV.

$L_1L_2L_3L_4$, lamps; $K_1K_2K_3K_4$, reaction keys; LC, contacts for lighting circuit; MC, contacts for marker circuit; $R_1R_2R_3R_4$, connections with automatic switchboard; F_1F_2, connections with marker; CS, copper strips on clock; H, clock hand connection; M, magnet; A, armature; R, ratchet engaging wheel; P, axle suspending arm, R–A; N, contacts; KY, kymograph; J, Jaquet time marker; Y, marker of right reactions; Z, marker of all reactions; B_1B_2, batteries.

not learn the order or any part of it, though it is repeated ten times in ten minutes.

The above arrangement of exposure and recording apparatus gives a very serviceable device for discrimination work. There

are no distractions. The eye may remain fixed on the stimuli. The reactions are simple and not fatiguing. The order of presentation is not repeated for sixty exposures and the right and wrong reactions are both recorded. Not the least attractive feature is the interest the reactions arouse in the subject, who willingly struggles with a long series "for the fun of it."

EXPERIMENTS IN BALL-TOSSING: THE SIGNIFICANCE OF LEARNING CURVES

BY JOSEPH PETERSON

University of Minnesota

Swift's curves of learning to keep "two balls going with one hand, catching and throwing one while the other is in the air," do not show the rapid initial decline characteristic of most motor learning curves. On the contrary they are in general "concave toward the vertical axis, which means, of course, that the progress was at first slow and then more rapid."[1] In Swift's curves vertical distances represent the average number of successful catches in ten trials, a day's practice; and the horizontal distances indicate the successive days, or the number of practice periods. By this method, as Thorndike has pointed out,[2] equal distances on the abscissa line do not mean equal amounts of practice. A rearrangement of the data of one of Swift's subjects so that the horizontal distances will stand for equal number of tosses results in a curve which up to 4,500 tosses may be roughly represented by a straight line.[3]

In animal and other experiments in which errors are significant, as is true also of ball-tossing, curves of errors have been given along, usually, with curves of other data. Just what data to take as the most significant in any learning process, is a question that is not settled, and one that must obviously be answered differently for different learning processes. In most cases, even when time, errors, and performance are plotted in separate curves, the data are liable to be misleading if taken to represent the entire learning that goes on. If, for instance, one of the curves shows a plateau

[1] Swift, E. J., 'Mind in the Making,' 1909, p. 177. The interpretation expressed in this quotation is erroneous, though it is very frequently expressed in psychological literature.

[2] Thorndike, E. L., 'The Psychology of Learning,' 1913, p. 121.

[3] Thorndike, *op. cit.*, p. 122.

we are too apt to conclude that learning was abated or wholly suspended at this point. This tendency is especially noticeable in cases of learning represented by a single curve.

The Experiment

The experiment here reported was devised to secure extensive data on ball-tossing in terms of errors, so that the results might be compared with those of other kinds of experimentation. Though other data—time and 'the feeling of efficiency' before practice—were obtained, the conditions of the experiment were not such as to make these data very reliable. They will therefore be considered only in a general way in the discussion.[1] The subjects in the experiments were sophomore, junior, and senior students taking their second half year's course in psychology. The experiment was performed in the spring of 1916. Each subject received written instructions as follows:

"Get two solid rubber balls, each of about 42 mm. diameter and weighing approximately 123 grams. Select for practice each day a period as free from interruption as possible, preferably in the morning before breakfast (or two periods daily, one in the morning and one in the afternoon). The place of practice, as well as the time, should be constant. A room with high ceiling and light walls is best; varying conditions of the sky would interfere with outside practice. Arrange something to catch, and thus to prevent from rolling away, the balls you miss. Practice the act of keeping two balls going with one hand, catching and throwing the one while the other is in the air. (This was illustrated to all subjects by the writer.) Count only the catches, beginning with that of the first ball tossed and ending with that preceding the first miss. This is one series. Each practice period is to continue until the total number of successful catches reaches 200, when it must stop. The two-hundredth, if stopped voluntarily, is not a miss. Rest between series whenever tired, but only for one minute. Continue, practicing regularly every day, until for five days you have made no misses.

"Keep a careful record of the number of catches in each series, the number and place in the practice period of rests, cause of miscatches when easily noticeable, time of day, conditions of light (never artificial), your own condition (and whether you feel that you are going to make a good record), chief difficulties and how overcome, total time of each practice period, and any other observations. Follow the directions closely, noting any departure from them. Be constantly on the alert for factors or conditions enabling you to explain any irregularities."

The instructions are slightly changed as published here, but only in unessentials, such as the rearrangement of sen-

[1] In general it may be said here that time curves seem to parallel rather closely error curves and that one's 'feeling of efficiency' before practice is, on the whole, not a very reliable index of efficiency as measured by actual results.

tences, and the making more explicit of detailed matters explained orally to all subjects together. All who found it possible were urged to take two periods of practice daily, of 200 catches each, so that we might have data on the distribution of practice. In such cases one practice was to take place in the morning and one in the afternoon, never both in the same half day. There were, of course, slight deviations in the sizes of the balls used by different individuals and slight irregularities in other details of practice. Occasionally, for instance, a day's practice would be omitted because of illness or absence, some of the students were obliged to practice throughout the experiment in artificial light, and occasionally one would forget in one's enthusiasm and run over the 200 limit of catches. All papers showing irregularities or carelessness judged sufficient to interfere with the reliability of results were excluded before the data for the writing of this report were collected. That is to say, the papers used in the report were *selected* only on the basis of regularity and carefulness of work, not on that of the kind of results secured! In all individual records used the details of the experimental procedure and results were sufficient, in connection with individual conferences and reports during and after the experiments, to enable the writer, the director of the experiments, to judge of their reliability. No records are included of individuals who had had previous practice in the performances, even though the subject had apparently lost the effects of the practice.

In all, the records of twenty-six individuals are included in the reports, though the data from three of these records are excluded from the averaged curves because the practice in each case was discontinued before a high degree of skill was acquired. Two of these were taken from the once-a-day group and one from the twice-a-day group. For reasons already explained the records of eight subjects were excluded from any consideration in this paper. The total number of persons performing the experiment was therefore thirty-four.[1]

[1] It should be stated also that the experimenter himself practiced at ball-tossing daily for three months in the spring of 1914. In the following year he obtained results

Results

The individual results of these experiments are given in the following tables and figures, so that the reader may see the extent of individual variations and that he may judge for himself on important points mentioned in the following discussion. It is important, moreover, to preserve these individual records, which in each case represent probably from ten to fifty hours' experimentation, so that they may be available for other investigators who may be doing work of a similar kind. Mere averages, we must also keep in mind, not infrequently cover up significant aspects of individual results, such as plateaus, initial 'spurts' and changes near the physiological limits of learning. In results such as these, deviations from central tendencies are not particularly illuminating, and have therefore not been given. They may be calculated by any one who desires them. It is more important for our purpose to give the individual results and the averages.

Explanatory Note.—In Table I. and Table II. a horizontal line between any two numbers in a column indicates one practice period missed; if more periods than one were missed the exact number is indicated by an italic figure at the right of the line. The columns bracketed at the beginning and end are not included in the averages given in Table III.

Table I. gives the results of the fifteen subjects who had one practice period each day. Table II. is the corresponding record of the group of thirteen subjects who had a practice period in the morning and one in the afternoon, or two in twenty-four hours. The first column, P. P., gives the number of the practice period. For example, 20 in the P. P. column in Table I. is the twentieth day of the practice and also the twentieth practice period, whereas in Table II. it represents the afternoon practice of the tenth day. The other columns give the number of errors—miscatches—made by the respective individuals in the practice periods of 200

from fifteen subjects, which agree substantially with those here reported. These results are not incorporated in the data of this report for the reason that the same uniformity of conditions of experimentation was not secured. The method used in the present grew out of the earlier work.

Table I

Individual Records of Once-a-day Group. Errors in 200 Catches

P. P.	Subjects														
	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*N*	*O*
1	209	104 —2	190	111	79	155	157	120	153	64	38	72	122	105	116
2	195	55	185	42	53	128	166	69	113	37	42	25	102	68	78
3	181	45	187	44	34	107	168	72	120	42	19	21	100	63	65
4	165	40	175	42	27	76	150	52	83	19	23	23	70	60	55
5	152	37	172	42	24	56	145	62	88	22	13	10	86	49	42
6	143	36	169	21	13	49	147	53	81	16	12	11	75	49	35
7	154	34	158	19	11	37	139	48	89	15	9	11	65	37	22
8	129	28	155	30	10	35	132	45	85 —2	12	10	6	62	32	26
9	116	26 —4	143	26	10	30	126	44	75	13	11	8	63	26	28 —3
10	119	14	141	24	9	26	111	44	78	11	15	7	57	19	16 —2
11	118	13	145	18	8	20	103	45	80	15	12	2	53	14	18
12	107	8	138	16	10	19	93	45	81	11	6	12	39	12	15
13	109	9	139	17	4	26	135	42	77	9	13	4	40	11	14
14	105	8 —4	135	15	9	20	77	37	77	12	11	3	37	11	14
15	112	9	121	16	8	20	71	34	77	10 —7	12 —2	2	38	9	17
16	103	5	112	9	8	15	55	36	80	11	5	6	44	10	15
17	108	12 —6	110	9	6	14	55	36	70	10	4	8	34	8	13
18	121	6	102	9	4	23	57	36	69	7	7	3	28	8	16
19	117	10	96	8	9	16	55	35	65	5	5	2	29	6	12
20	107	10	96	7	9	14	54	37	79	7	3	3	20	7	7
21	103	10	96	6	5	15	55	37	68	7	4	3	27	7	6
22	104	10	85	6	5	12	53	32	69 —2	5	6	3	20	6	5
23	106	6	79	9	5	11	53	25	66	6	4	1	26	5	5
24	109	9	65	9	5	11	50	24	63	5	3	3	33	4	6
25	103	10	63	10	6	20	55	24	64	5	2	5	21	4	5
26	101	7	67	7	8	18	51	27	66	5	4	3	23	4	4
27	106	10	57	6	4	14	50	26	53	4	3	2	18	4	3
28	109	8	59	7	4	15	45	24	58	4	3	3	14	4	8
29	102	8 —6	45	9	5	13	77	24	45	4	2	3	14	4	4
30	100	5	43	2	4	13	55	23	46	4	2	10	18	4	5
31	105	5	*	8	5	10	47	22	44	3	0	5	13	11	3
32	108	5		14	6	9	39	23	40	3	0	4	11	11	5
33	104	5 —2		7	5	19	46	18	36	3	2	3	12	8	2
34	99	2 —2		9	5	16	35	15	44	4	2 —2	2	11	7	3

P. P.	Subjects														
	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*N*	*O*
35	107	7		3	6	10	36	18	33	4	2	3	8	7	2
36	102	2		4	4	9	33	12	34	2	0	4	7	7	3
37	106	3		7	4	9	31	13	31	2	0	3	8	7	0
38	97	1		5	4	9	30	14	37	3	0	3	16	6	1
								—3							
39	99	*		7	5	8	27	13	28	2	0	4	7	6	1
40	101			7	7	8	33	12	28	2	0	2	12	6	1
41	100			5	4	8	30	12	32	2	*	3	6	5	1
42	104			5	5	10	30	7	29	3		4	10	5	0
43	102			6	3	9	39	8	26	2		3	8	4	1
44	*			7	5	9	34	4	28	2		5	12	3	1
45				3	4	8	29	8	24	2		2	8	3	0
46				5	4	10	24	5	25	2		4	7	3	2
47				5	2	10	23	6	24	2		10	9	3	2
48				9	2	9	22	6	22	2		5	7	3	1
49				5	3	7	27	7	20	2		2	8	2	1
50				4	4	5	35	*	18	2		2	5	2	0
51				4	2	*	30		24	2		2	8	2	0
52				4	3		26		17	2		3	5	2	0
53				5	3		21		20	0		3	11	2	*
54				2	0		18		18	0		0	6	2	
55				2	1		12		*	2		2	9	2	
										—5					
56				3	1		9			2		1	4	1	
57				3	1		11			0		2	4	1	
58				1	1		7			0		1	5	1	
59				1	1		4			0		2	8	1	
60				3	2		3			0		4	4	*	
61				1	2		*			0		4	3		
62				4	2					*		1	4		
63				1	3							2	5		
64				1	1							2	6		
65				3	*							2	4		

(*D*'s record for successive days continues as follows: 3, 2, 2, 0, 1, 1, 0, *; *L*'s 0, 2, 2, 2, 2, 2, 1, *; *M*'s 5, 4, 3, 6, 6, 5, 4, 8, 4, 2, 3, 0, 4, 5, *.)

successful catches each. Horizontal lines between any two numbers in a given column indicate an interruption at that point in the practice, due to one of several conditions, such as temporary illness, absence from the place of practice, etc. A small number at the right of one of these lines shows the number of periods omitted. When no number is given with the line only one period was missed. Thus the record of '*B*' shows two periods missed between the ninth and the tenth, one period between the sixteenth and the seventeenth, and so on. This student has the poorest record so far as regu-

Table II

Individual Records of Twice-a-day Group. Errors in 200 Catches

P. P.	Subjects												
	P	*Q*	*R*	*S*	*T*	*U*	*V*	*W*	*X*	*Y*	*Z*	*AA*	*BB*
1	103	161	98	47	91	148	200	66	72	117	87	174	213
2	79	143	95	34	74	113	178	52	62	58	80	146	205
3	66	112	93	20	69	110	137	46	26	57	38	109	208
4	62	111	91	9	46	111	121	38	25	76	27	96	200
5	52	78	90	7	33	97	123	33	17	24	30	70	203
6	48	68	87	7	26	79	121	55	16	23	29	61	200
7	48	66	85	4	18	86	109	28	19	19	15	44	195
8	37	55	84	5	26	108	109	21	12	13	17	40	200
9	34	57	83	5	17	79	121	18	9	12	15	37	180
10	34	59	80	2	17	88	91	17	13	10	18	33	187
11	28	36	78	4	17	65	102	15	8	8	12	28	185
12	31	29	75	3	12	73	77	13	8	10	13	28	165
13	37	29	74	3	13	71	102	12	9	9	9	27	163
14	34	42	73	3	11	53	71	12	6	6	11	25	170
15	32	30	71	1	10	50	109	9	6	5	11	26	173
16	26	32	69	1	12	40	82	8	8	6	11	27	168
17	27	22	65	3	11	42	81	9	5	5	12	27	150
18	31	28	60	1	7	39	71	7	10	4	11	30	153
19	20	23	58	1	12	27	74	6	6	4	10	41	149
20	22	26	55	2	12	25	80	9	8	8	9	32	154
21	22	21	51	1	8	33	69	7	4	4	8	30	152
22	21	25	49	1	11	35	65	6	6	4	8	25	135
23	16	18	48	1	5	27	48	5	6	4	5	24	151
24	25	15	43	1	6	25	44	5	11	7	6	20	140
25	21	21	40	1	6	33	36	3	6	3	4	20	147
26	14	18	37	1	6	29	36	3	7	6	4	19	138
27	14	20	30	0	6	25	50	2	3	3	5	20	120
28	13	15	26	0	7	28	64	2	4	6	5	20	127
29	14	15	25	0	4	22	53	2	3	2	5	18	100
30	16	18	21	0	7	22	50	1	3	4	4	15	105
31	9	12	35	0	8	23	55	1	3	3	5	17	109
32	10	11	29	*	7	21	40	3	3	3	5	16	105
33	8	17	24		7	24	54	1	4	3	2	18	95
34	8	18	21		4	22	37	2	3	2	5	21	92
35	10	18	19		7	21	41	2	2	5	4	12	99
36	5	18	17		3	23	63	2	3	3	7	9	85
37	15	13	15		6	21	46	2	4	3	5	8	94
38	7	15	19		1	25	33	3	2	3	6	7	110
39	10	13	16		4	27	37	3	2	1	2	9	80
40	3	12	14		4	20	37	2	2	2	2	16	84

P. P.	Subjects												
	P	*Q*	*R*	*S*	*T*	*U*	*V*	*W*	*X*	*Y*	*Z*	*AA*	*BB*
41	7	12	10		5	14	53	1	2	2	0	17	70
42	5	12	9		3	19	39	1	1	2	4	9	72
43	4	13	12		5	21	41	1	2	1	3	8	76
44	3	11	23		5	17	42	1	0	2	1	5	74
45	4	12	20		6	13	31	1	3	1	0	5	78
46	4	10	17		2	14	33	1	1	1	1	6	65
47	8	9	14 —2		3	8	37	1	1	1	0	10	72
48	3	11	13		3	10	34	3	1	2	0	4	74
49	6	12	11		5	10	27	1	1	2	2	4	65
50	6	14	9		6	13	38	3	0	1	1	6	60
51	2	10	9		5	14	33	1	0	2	3	4	67
52	3	10	8		4	10	30	1	0	1	0	3	69
53	3	7	12		4	12	32	2	1	2	0	2	65
54	5	4	10		1	12	28	1	0	1	2	3	74
55	5	9	9		4	10	35	1	0	2	0	5	72
56	4	6	8		2	7	31	2	0	1	1	3	55
57	3	9	7		0	5	35	1	0	1	1	2	55
58	2	7	7		4	10	34	0	0	1	0	0	57
59	6	5	8		2	8	21	1	*	1	0	3	65
60	3	8	7		2	8	25	1		1	0	2	75
61	5	9	7		2	12	31	1		1	0	0	87
62	4	4	7		4	12	15	0		1	0	0	70
63	4	6	6		3	11	23	0		1	*	2	72
64	3	6	6		4	5	35	0		0		*	66
65	3	5	6		3	7	23	1		0			60
66	3	4	6		4	7	25	1		0			71
67	5	4	7		2	2	24	2		0			57
68	2	2	6		1	6	19	0		0			55
69	5	5	7		5	9	24	0		*			64
70	3	1	6		3	8	28	1					67
71	4	3	4		2	12	16	*					62 —2
72	3	5	5		0	9	22						62 —2
73	4	2	5		1	9	15						72
74	2	5	5		3	5	16						57
75	2	5	5		2	11	19						65 —2
76	3	5	6		2	6	23						55 —2
77	2	2	6		0	8	21						68
78	5	4	4		0	6	18 —5						65
79	5	4	4		0	12	16						70

P. P.	Subjects												
	P	*Q*	*R*	*S*	*T*	*U*	*V*	*W*	*X*	*Y*	*Z*	*AA*	*BB*
80	2	2	4		0	8	16						64
81	5	6	5		0	13	16						60
82	2	7	4		*	4	23						62
83	5	3	4			4	30						58
		—*3*											
84	5	4	3			1	15						*
	—*2*												
85	4	3	4			5	15						
86	3	4	3			8	10						
							—*3*						
87	3	5	3			3	27						
88	1	5	4			6	12						
89	1	1	4			4	24						
90	4	7	4			3	17						
91	2	6	3			4	13						
92	5	4	3			3	20						
93	3	3	3			5	20						
94	3	4	3			4	12						
95	1	3	2			5	10						
96	1	1	3			5	9						
97	2	3	2			6	11						
98	2	1	2			7	10						
99	1	2	*			1	15						
100	3	0				6	11						

(The four records incomplete at this point go as follows: *P*, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, *; *Q*, 0, 0, 5, 1, 3, 2, 2, 0, 0, 0, 6, 2, 3, 1, 2, 0, 3, 1, 2, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, *; *U*, 3, 3, 4, 6, 2, 3, 2, 2, 3, 2, 3, 4, 6, 2, 7, 3, 4, 3, 4, 6, 6, 3, 4, 4, 3, 4, 4, 1, 3, 3, 2, 1, 1, 1, *; and *V*, 13, 15, 8, 8, 9, 10, 13, 4, 8, 15, 5, 10, 8, 11, 5, 7, 6, 13, 5, 6, 6, 8, 6, 7, 3, 6, 7, 6, 6, 11, 4, 12, 5, 4, *. Two of these, it will be noted, continue to the 134th period of practice, and are discontinued still unlearned, or below the standard set.)

larity of practice is concerned. The results of these irregularities are not obvious, though they no doubt affect the validity of any conclusions that may be drawn from the averages of the two groups as to the relative values of one practice a day and two practices a day.

Figs. 1 and 2 show the individual learning curves. Each curve is to be identified by the letter corresponding to the column in the tables which it represents. In these curves vertical distances represent the number of errors, or miscatches, and the horizontal distances are successive practice periods of two hundred catches each. It is important to

note the frequent absence of plateaus in those curves which are carried to, or nearly to, a point of zero errors in 200 catches. Too much must not be made of this fact, however, as we shall see when we come to the interpretation of these, and of other, learning curves. The curves are marked by rapid initial declines, having the general form of the so-called normal curve of motor learning. *Up to the present time this rapid initial decline in errors or in time has been interpreted as showing a corresponding rapid initial learning—an unfortunate error*, due largely, it appears, to the ease with which 'errors' and time in seconds are measured, and to the lack of thoroughgoing criticisms of the significance of any such arbitrary measurements of learning. Various ingenious reasons why there should be rapid initial learning corresponding to the rapid initial decline of the error, or time, curves have been 'thought out.' We shall later on consider this matter more fully.

Cross lines on the curves, and the numbers that occasionally occur with them, indicate practice periods omitted, as has already been explained in connection with the tables.

Many interesting individual differences are noticeable in Figs. 1 and 2. *A* was a student of high rank in the class work. Her curve seems to indicate that some peculiar, obstructing habit developed early in the experiment, bringing the curve almost to a plateau. It is impossible to say how the curve would have developed if the experimentation had continued to the point of completion. In such a case a little help by the experimenter at the point of discontinuance would doubtless have resulted in a rapid drop at this point. Without such help *A* seemed to have approached a sort of pseudo-physiological limit, the learning likely being limited or obstructed by the formation of habits for certain constant errors. Curve *I* is a very peculiar one; it is also the curve of a woman. This student was older than most of the others, was above the normal in class standing, but was evidently less adaptable in this kind of physical performance than most of the other students. *C* and *G* are also women making poor ball-tossing records, *C* making also only a passing record

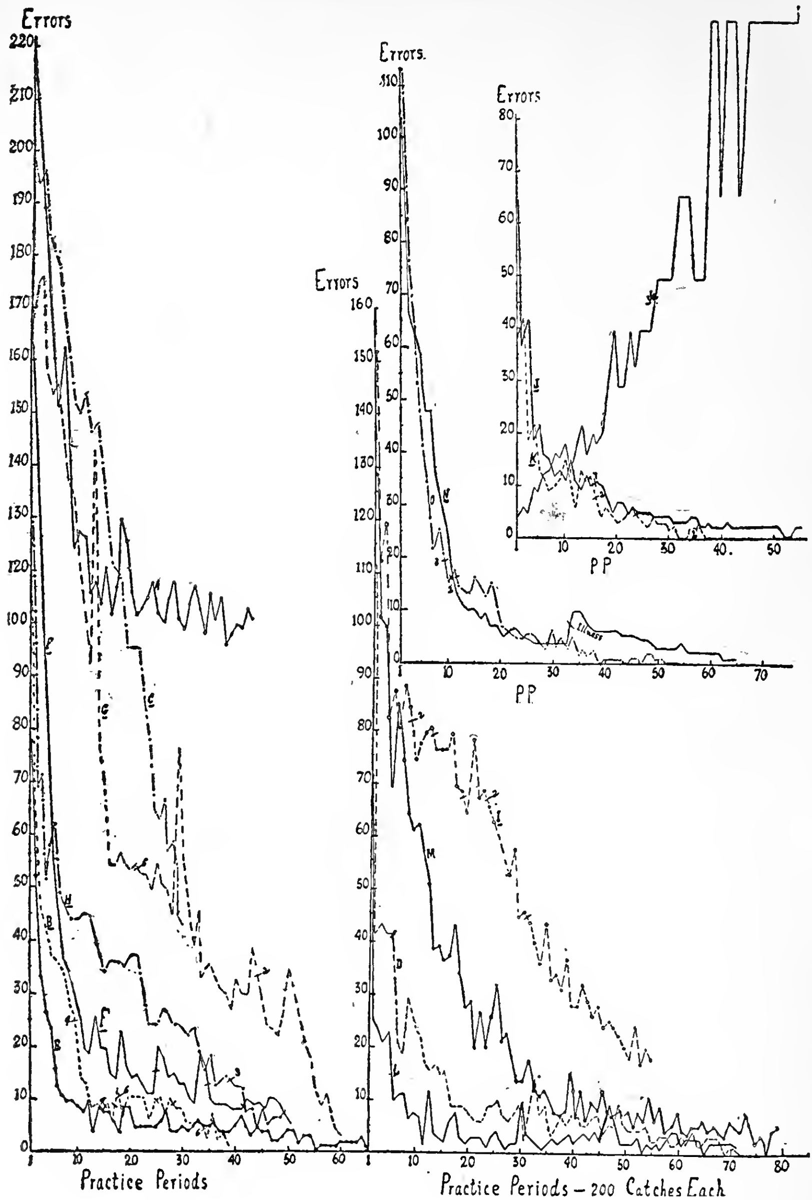


FIG. 1. Individual error curves of fifteen subjects practicing once a day. The small cross-lines and figures indicate practice periods missed. *Ja* is the curve of *J*'s average number of catches per practice period. The learning of any other subject may be given in a similar way. Letters underscored indicate women.

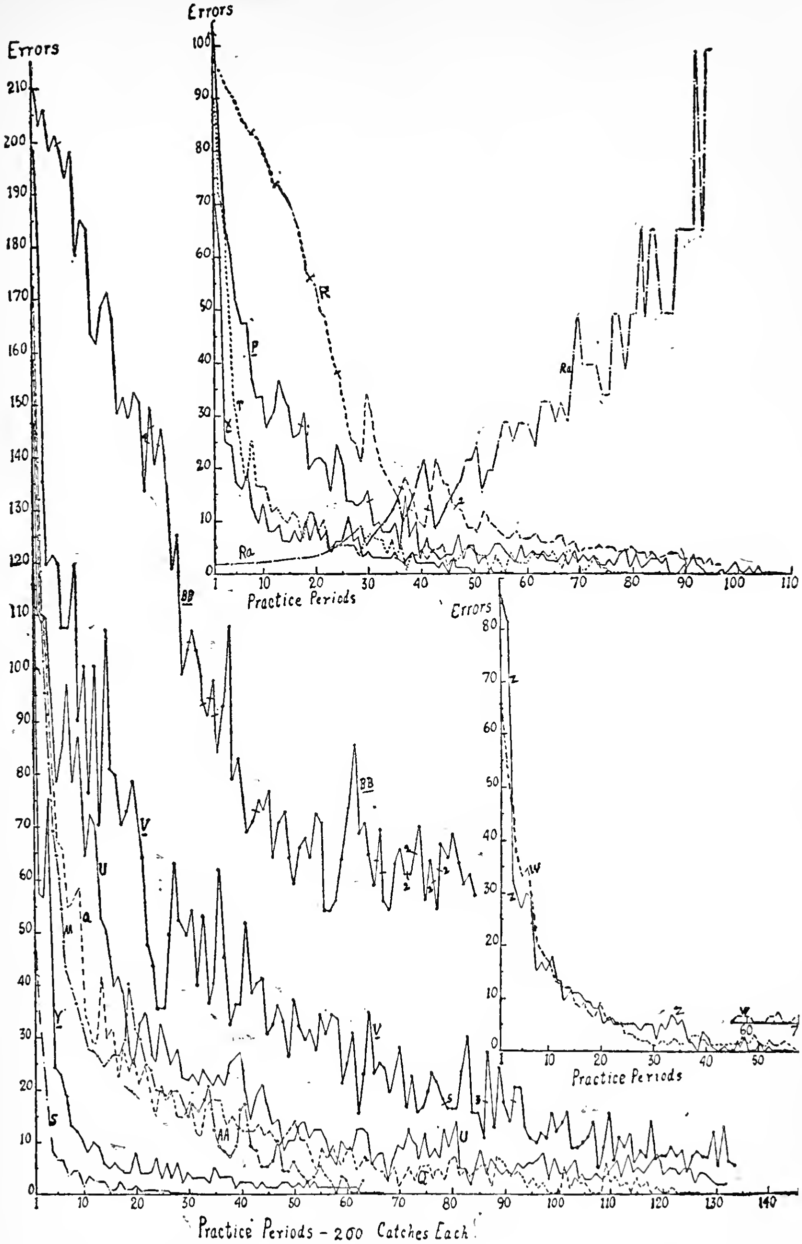


FIG. 2. Individual error curves of thirteen subjects practicing twice daily. Omissions of practice periods are indicated as in Fig. 1. *Ra* represents *R*'s learning in terms of average number of catches per practice period. Letters underscored ndicate women.

in the psychology class. That women do not make the poorest records invariably is evident from the curves of *B*, *E*, *J* and *K*, all women. In Fig. 2 the three students making the poorest progress in the experiment are *V*, *BB*, and *R*, the first two being women and the last a young man who received a D grade in the psychology course. Here again *V* was one of the very best students in two sections of general psychology taught by the writer. *Y* is a woman of C grade in her class, and *S*, making the best record of all in the experiment, is a man who has had considerable experience playing ball. From these records it would be hard to give any reliable information for prediction as to one's success at the performance here practiced. Early interest in ball playing and practice at throwing and catching balls as children do in ball games would seem to be one of the best indicators of success in this experiment, but even here we could not say whether the practice had helped or whether the particular nerve and muscle conformation inherited had predisposed the subject both to interest in ball playing and to ability in ball juggling. The latter would seem to be the more probable condition.

Curves *Ja* and *Ra* should be studied carefully in relation to their reciprocals *J* and *R*.[1] It will be noted that, in conformity with the view suggested in this paper, the error curve changes most rapidly in the early part of the learning where the average attainment is small, or where the series are short. Each series is necessarily ended by an error. In the initial learning, then, there are a great number of series, or different trials, and hence a very small change in each, on the average, makes a big change in the total errors. The error curve is therefore liable to fluctuate more in the initial phases with the average curve advancing more constantly, while in the final part of the learning the error curve will assume a more constant or smooth appearance while the average curve will fluctuate between wide limits. If one to

[1] Note that the term 'reciprocal' is used here not in the strict mathematical sense of values whose product is equal to unity. Their product, with a slight qualification developed in subsequent pages, is equal to a number representing *K*, the constant of practice.

two hundred catches can be made successfully in a series, the total errors for two hundred catches will not change more than one or two units in successive practice periods, while the average number of catches may vary from less than fifty up to two hundred absolute units.

This reciprocal relation obtains between errors and average number of catches because the former directly condition the latter and *vice versa*.[1] This state of affairs is not so obviously present in some types of learning processes, but as we shall see it is well worth enquiring into the relation of errors to attainment in most such experiments.

Table III. gives the averages of the individual errors, column A. E. For reasons already explained these averages do not include the records of subjects *A*, *C*, and *BB*, bracketed at the top and the bottom of the columns. In determining these averages it has been necessary to take note of the fact that some of the columns terminate before reaching zero, no errors. In such cases the averages have been weighted as follows. If, for example, ten columns were being averaged, and one of these discontinued with the 50th practice period (P. P.), it would leave only nine columns from this point. Suppose the last five records of the column terminating here were 8, 5, 6, 6, 7. Suppose also that the sum of the remaining nine records for the fifty-first practice period was 27, then it would seem fair to add to the average, 3, of the nine columns about .3 $\left(i.\ e., \frac{27 + 6}{10} = 3.3\right)$. The amount added would be gradually decreased to represent the decrease that the discontinued record would probably have had if it had been continued. This weighting of averages is to prevent the appearance in the curve of a sudden change that might be misinterpreted. None of the averages in the twice-a-day group is weighted. Only the last seven averages in the once-a-day group are weighted; these are of course less

[1] An error in ball-tossing ends in fact one *trial* (termed in this paper a series); and it is obvious that the number of trials *times* the average number of catches per trial *equals* the total number of catches. Hence dividing the number of trials ('errors') into the total number of catches we get the average number of catches, 'average attainment.' See next page.

Table III

Averages of Errors in 200 Catches and Average Number of Successive Catches[1]

Once-a-day Group						Twice-a-day Group					
P. P.	A. E.	A. C.	P. P.	A. E.	A. C.	P. P.	A. E.	A. C.	P. P.	A. E.	A. C.
1	107	1.8	34	12	16.7	1	97	2.1	34	12	16.7
2	75	2.6	35	11	18.2	2	93	2.2	35	12	16.7
3	69	2.9	36	9.3	21.5	3	74	2.7	36	13	15.5
4	55	3.6	37	9.1	21.9	4	68	2.9	37	12	16.7
5	52	3.9	38	10	20.0	5	55	3.6	38	10	20.0
6	46	4.3	39	8.3	24.1	6	50	4.0	39	10	20.0
7	41	4.9	40	9.1	21.5	7	45	4.4	40	10	20.0
8	39	5.1	41	8.3	24.1	8	44	4.5	41	10	20.0
9	37	5.4	42	8.3	24.1	9	41	4.9	42	9.7	20.9
10	33	6.1	43	8.4	23.8	10	38	5.2	43	9.3	21.5
11	31	6.5	44	8.5	23.5	11	33	6.1	44	9.2	21.8
12	28	7.2	45	7.0	28.6	12	31	6.5	45	8.0	25.0
13	31	6.5	46	7.2	27.8	13	33	6.1	46	7.5	26.7
14	26	7.7	47	7.5	26.7	14	29	6.9	47	7.5	26.7
15	25	8.0	48	6.8	29.4	15	30	6.7	48	7.0	28.6
16	23	8.7	49	6.4[2]	32.8	16	27	7.6	49	6.8	29.4
17	21	9.5	50	6.4	32.8	17	26	7.7	50	8.1	24.7
18	21	9.5	51	6.7	29.9	18	25	8.0	51	6.9	28.9
19	19	10.5	52	5.9	33.9	19	24	8.3	52	6.6	30.3
20	20	10.0	53	5.6	35.7	20	24	8.3	53	6.4	32.8
21	19	10.5	54	4.2	47.6	21	22	9.1	54	5.6	35.7
22	18	11.1	55	4.2	47.6	22	21	9.5	55	6.7	29.9
23	17	11.8	56	3.3	60.6	23	17	11.8	56	5.4	37.0
24	17	11.8				24	17	11.8	57	5.3	37.7
25	18	11.1				25	16	12.5	58	5.4	37.0
26	17	11.8				26	15	13.3	59	4.6	43.5
27	15	13.5				27	15	13.3	60	4.8	41.7
28	15	13.5				28	16	12.5	61	5.7	35.1
29	16	12.5				29	14	14.3	62	3.9	51.3
30	15	13.5				30	13	15.5	63	4.8	41.7
31	14	14.3				31	14	14.3	64	4.9	40.6
32	13	15.5				32	12	16.7	65	4.0	50.0
33	13	15.5				33	14	14.3			

reliable than the others. The averages of the errors in this group are necessarily discontinued earlier than those of the twice a day group; for a given number of days' practice will make the column only half as long as with two periods of practice daily. A number of the subjects did not get started early enough in the semester to complete the learning by the close of the academic year. While this is regrettable, it is to be noted that the records are complete enough for reliable data on the most important matters raised in the following

[1] P. P. = practice period; A. E. = average number of errors in 200 catches; A. C. = average number of successive catches in the P. P.

[2] Averages of this column are 'weighted' below this point, as explained in the text.

discussion of the significance of the curves. Our experiment is only to be regarded as preliminary, so far as affording data for finer points on learning curves is concerned. More data are desirable particularly toward the end of the learning process in the neighborhood of the 'physiological limit.' It will be noted that so far as our averages go in the once-a-day practices, they show more rapid learning for equal practice than do those of the other group with more frequent periods. This agrees with most other experiments on the distribution of practice. Individual variations, however, are so great and the number of subjects so small as to invalidate any conclusion on this point from our data. An unequal proportion of each sex in the two groups is another factor that may be responsible in a measure for this result, though the probability seems to be that a more equal sex distribution would have made the once-a-day results even comparatively better than they were.

In Table III. are also given the average number of successive catches, A. C., in each practice period, the average number of catches without an error. These averages were obtained simply by dividing the average number of errors, A. E., into 200 catches. In order that the successive practice periods might be as nearly equal as possible, the subjects were asked always to stop any series voluntarily, if it did not end there by a miscatch, when the total number of catches of the practice period reached 200.

But this voluntary discontinuance of the final series at 200 catches makes an error in our results which seemed unavoidable so far as the experimental procedure is concerned. This error is fortunately negligible in the first part of the learning process, that with which we are most concerned in this study, since no subject could have exceeded very far the 200 limit by continuing the series voluntarily stopped. As the average number of catches increases this error in our results becomes larger. More specifically, if a subject averaging four catches per series ends by a miss, say, at 198 catches and 50 miscatches or errors, the chances are that he could on the next series reach 202 catches (with 51 errors). Volun-

tarily stopping at the 200th catch therefore gives him fewer errors than he should have, and correspondingly increases the average number of catches. In this case one half error should have been added to those actually made. Instead of having an average of 4 catches to the series (*i. e.*, 200 ÷ 50) he should really have 3.96 (*i. e.*, 200 ÷ 50.5). If, on the other hand, one averaging nearly 50 catches to the series misses, say, at a total of 175 catches with 4 errors, and then makes up the 200 total and stops voluntarily, he should have not an average of 50 catches to the series (200 ÷ 4), but only one of 44.44 catches (200 ÷ 4.5). In the former case the voluntary discontinuance at 200 made a negligible error in results (1 per cent.), while in the latter the miscatches, 'errors,' should be 12.5 per cent. more than would appear from our method of procedure, and the average number of catches would be 12.5 per cent. less.

Fig. 3 presents in a graphic form the data of Table III., without any correction for the error just pointed out. Curves *E*, I., and *E*, II., represent the number of errors (miscatches) in the 200 catches of the one-practice-a-day group and the two-practices-a-day group, respectively. One practice a day seems to be a slightly more effective distribution than two a day. It is important to note—and this is one of the main contributions of this paper—that these curves are the so-called *normal curves of motor learning*, showing a rapid initial decline and a more gradual decline later, as is the rule with such curves. That this is not evidence of rapid initial learning is obvious from the fact that the other curves in this figure, *A*, I., and *A*, II., representing the average number of successive catches without a miss, have a different form altogether though they represent the same learning process as the *E* curves. That is to say, the progress of the one-practice-a-day group is represented by the curve *E*, I., and also by the curve *A*, I. If we consider only the former we are in danger of concluding—as has been the practice among practically all psychologists for error curves up to this writing—that the learning begins at a rapid rate and gradually decreases in absolute amount in the successive practice

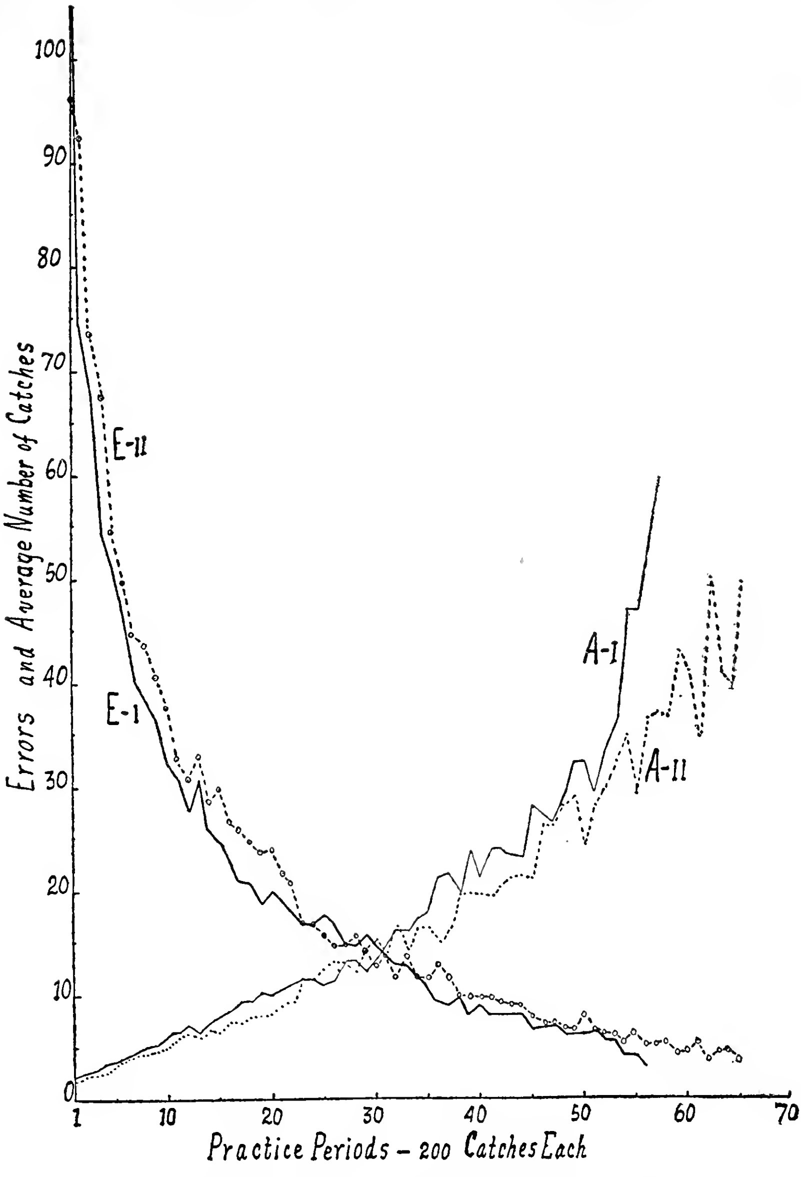


FIG. 3. *E*, I. and *A*, I. are the error curve and the curve of average number of catches per practice period of thirteen subjects practicing once a day. *E*, II. and *A*, II. are the corresponding curves of twelve subjects practicing twice daily.

periods, being very slow toward the end of the experiment. On the other hand, if we consider only the latter curve, just the reverse of this condition will seem to be true. The only fair conclusion, taking both curves into consideration, would seem to be, so far as the data considered are concerned, *that we do not know anything reliable at all concerning the kind of learning in question, as to whether it is initially more rapid or more slow than at certain later stages.* It would seem safest to assume that in the case of ball-tossing, at least, learning goes on throughout, up to a point approaching a physiological limit, at about a constant rate. As we shall see later, there is good reason to suspect that the same thing is true of other learning processes. There is likely no real difference in this regard between the so-called motor learning and the sensory discrimination learning in animals. The distinction usually made between these 'kinds of learning' has long seemed artificial and erroneous to the writer, but only recently, in working up the results of the present experiment, did he secure a good basis in fact for the view that the distinction is incorrect. Every case of motor learning must involve some sort of 'sensory discrimination,' and conversely, all sensory stimuli must come to have some sort of motor significance.

One remarkable result of our experiments is the almost complete absence, in the individual as well as in the averaged results, of any important plateaus. As we shall see later, this may not mean that there were no actual plateaus in the learning, but it may indicate that certain kinds of curves tend to conceal plateaus. The very fact, however, that some kinds of curves tend to show plateaus where there are no plateaus by a different statement of results, or at least no marked plateaus,[1] may tend to cast doubt on the whole doctrine of plateaus in learning.

[1] A degree of evidence for this statement comes out in subsequent pages. Plateaus are shown best in the error or the time curves if they occur early in the learning process; if they occur in the latter part of the learning they show best in the average attainment curves, but are practically concealed in the error curves. See Fig. 10, for example. Some methods of plotting show them much better than others. If, moreover, the average attainment curve (which in ball-tossing seems to be concave to the y-axis) has a rapid initial rise because of actual rapid initial *learning* a plateau may seem to

An experiment of the kind here reported, taking a considerable amount of time daily from each of a large number of subjects, makes it impossible, under the conditions of full-time teaching by the experimenter, to keep as close check on each student as is desirable. Even if the experimenter were entirely free from other duties it would be very difficult to get the time of a great number of subjects regularly for the number of months that this experiment requires. Closer personal supervision than was possible to the experimenter in the present case might do something toward reducing individual differences, and especially toward explaining the causes of certain marked failures to learn.

In the introspections, observations, and explanations of the irregularities many interesting points respecting difficulties and the methods of their solution, or failure of solution, came out. These are perhaps not different from those already reported in previous studies, and need not detain us here. One of the very best psychology students of the subjects failed utterly to get a fair start, even with a little personal assistance—unfortunately somewhat too late—from the experimenter. She had had almost no experience in throwing and catching balls, and could not, after the day she came for special help, break up in the time available a confusion of impulses which blocked one another much as in the attempts of a stutterer to talk. Frequently she could not even throw up the first ball and hold the other one, but would throw both together in a spasmodic manner making them go in different directions. After a few weeks of practice she was seldom able even to catch the first ball tossed. The effort was then discontinued, and her results are not included in the data of our tables.

Another subject, *A*, whose results are not included in the averages because they did not reach a point of even fair efficiency, advanced rather slowly after a short initial gain, and at the point of discontinuance the process of learning

occur between this rapid rise in the curve and the one that occurs in the latter part of the curve, due purely to the *mathematical relations* of errors to average attainment. (See p. 199.) Such plateaus are not *plateaus in learning* at all. Just how marked a plateau is, is not easy to determine absolutely.

came to a condition approaching a plateau. This may have been due to the development of some such confusing tendencies as referred to in an extreme form in the preceding paragraph. This subject did not get time to complete the process up to a point that would indicate the real nature of this difficulty. The curve in question is given in Fig. 1. It should be noted that the learning had not come to a complete standstill. *A* was also a student making a good grade in the psychology course; she had had but little experience at ball playing in childhood. Subject *V* is another woman of whom the same things may be said, but her record was carried out to the 134th practice period (26,800 catches!), at which point she still made four errors in the 200 catches. All those who made extremely bad records were women, though some of the women subjects did exceptionally well. On the other hand, some of the men who had had extensive practice at playing ball advanced at an exceedingly rapid rate. Subject *S* is the best example in our groups. On the whole it appears that men are more rapid learners in ball-tossing of the kind practiced in the present experiment than are women, though our data are not sufficient, in view of the degree of individual variation found, to justify a detailed comparison or any definite conclusion on the matter. For the aid of any one who may desire to test this aspect of the problem out further and who may find use of our own data it may be added that the following subjects were women: *A*, *B*, *C*, *F*, *G*, *H*, *I*, *J*, *K*, *N*, *P*, *V*, *X*, *Y*, *BB*. Since the results averaged in the once-a-day group are of five men and eight women, and those averaged in the twice-a-day group are of eight men and four women, it is possible that the superiority of the former group would be greater with a more equal distribution of men and women. Since the time for practice had to be determined by the convenience of the subjects a lack of balance of this kind was not easy to avoid. It may be suggested here, too, that the very poor records of some of the women may not be due simply to the lack of practice at ball playing in childhood; on the contrary the failure to take up with this popular kind of sport may have been due

to an innate unfitness for it which likewise unfits the subject for ball-tossing.

The reports of the subjects show various individual differences in procedure, such as differences in methods by which the balls are tossed, in the methods of rotating them in the air to prevent collisions, in the distribution and frequency of 'bad days,' in the degree of correlation between feeling of efficiency and actual efficiency as shown by results (this correlation on the whole being very low), and in various other respects. A complete analysis of these differences with the data at hand would contribute but little to previous work, and is somewhat aside from our present purpose.

Statements like the following from *D*, which statement, by the way, came spontaneously from a number of the subjects, show the value of extensive experiments of this kind for students being initiated into scientific psychology. *D* writes: "I now feel as if I should like to continue something of this kind and gradually become proficient in several things. This experiment has surely taken the place of several books so far as learning is concerned, and I am very glad that I have had the opportunity to spend whatever time I have spent on it." The apparatus is simple, the directions are not hard to follow, and the working up of his own results is a most instructive exercise to the student.

We recur now to the error in results pointed out above.[1] Correcting for the effects of voluntarily stopping at the 200th catch would make the average attainment curves—those representing the average number of catches—appreciably less concave to the y-axis than is shown in Fig. 3; *i. e.*, more nearly approaching a straight line. In the averaged results of all the subjects it would make a difference of .5 per cent. in the first practice period; of nearly 1.5 per cent. in the tenth; of about 2 per cent. in the twentieth; of about 3.5 per cent. in the thirtieth; of about 5 per cent. in the fortieth; of about 7 per cent. in the fiftieth; and of nearly 10 per cent. in the fifty-sixth. In general it would seem fair to assume that the last series, which is stopped voluntarily each time,

[1] *Supra*, p. 197.

varies at random all the way from zero catches to the number of successive catches possible at any given stage of learning in question. We ought therefore to count, on the average, one half error for this series voluntarily stopped. Adding $\frac{1}{2}$ to the average of the recorded miscatches (errors) and dividing the resulting number into 200, we get the corrected average number of catches per series. A corrected table of the averages of the two groups of twenty-five subjects given separately in Table III. is given below (Table IV.). The formula used for determining the values A.C. in this table is

$$\frac{K}{e+\frac{1}{2}} = a.$$

This is a better empirical formula for conversion than that given above.

TABLE IV

A TABLE OF CORRECTED VALUES OF a AND e; RECORDS OF ALL SUBJECTS OF BOTH GROUPS ARE AVERAGED TOGETHER (28 SUBJECTS)

P. P.	E.	A. C.	P. P.	E.	A. C.
1	102.5	2.0	29	15.5	13
2	84.5	2.3	30	14.5	14
3	72.0	2.8	31	14.5	14
4	62.0	3.2	32	13.0	15
5	54.0	3.7	33	14.0	14
6	48.5	4.1	34	12.5	16
7	43.5	4.6	35	12.0	17
8	42.0	5.0	36	11.6	17
9	39.5	5.1	37	11.0	18
10	36.0	5.6	38	10.5	19
11	32.5	6.2	39	9.6	21
12	30.0	6.7	40	10.0	20
13	32.5	6.2	41	9.7	21
14	28.0	7.2	42	9.5	21
15	28.0	7.2	43	9.3	21
16	25.5	7.8	44	9.3	21
17	24.0	8.4	45	8.0	25
18	23.5	8.5	46	7.9	25
19	22.0	9.1	47	8.0	25
20	22.5	8.9	48	7.4	27
21	21.0	9.6	49	7.1	29
22	20.0	10	50	7.8	26
23	17.5	11	51	7.3	28
24	17.5	11	52	6.8	30
25	17.5	11	53	6.5	31
26	16.5	12	54	5.4	38
27	15.5	13	55	6.0	34
28	16.0	13	56	4.9	41

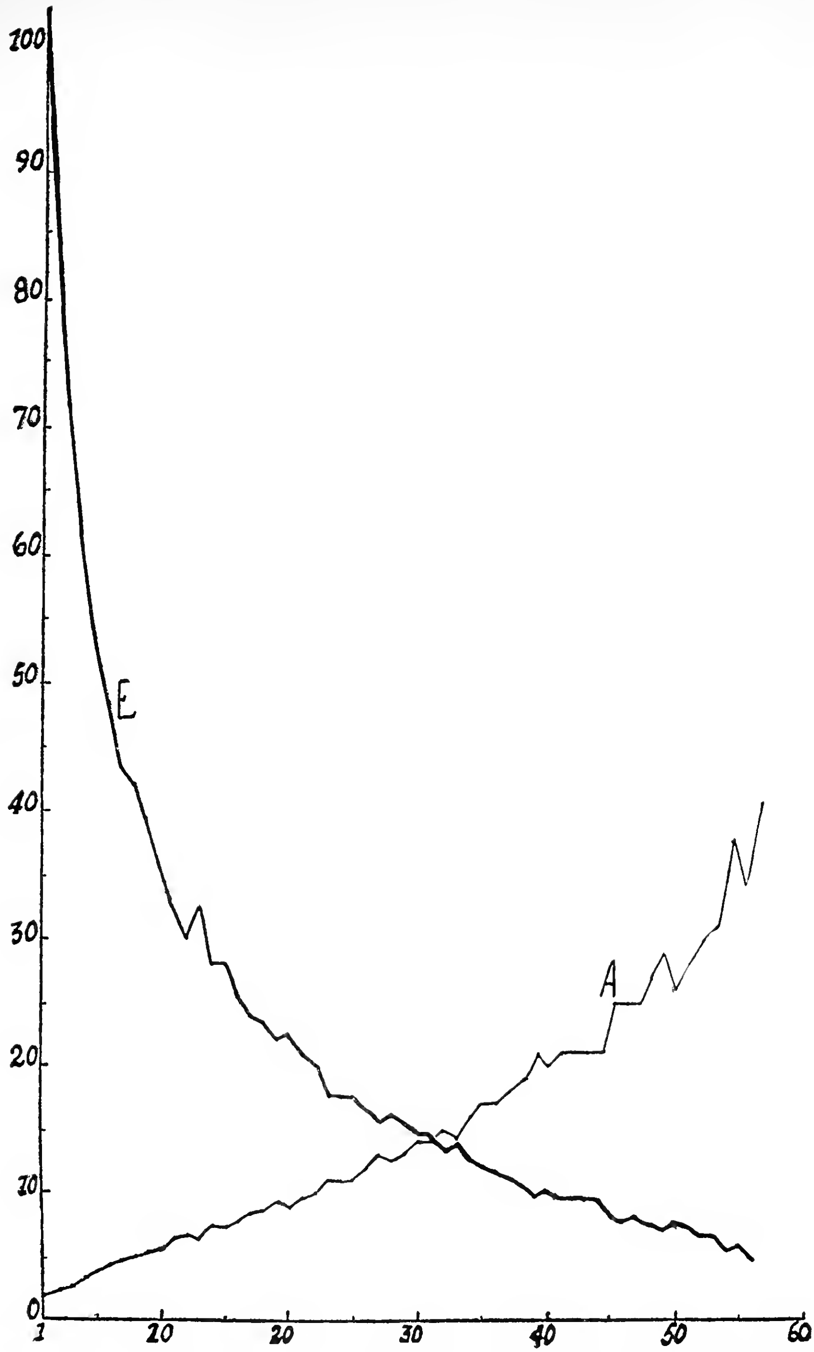


FIG. 4. Corrected error and average-number-of-catches curves of twenty-five subjects.

Fig. 4, giving graphically the data of Table IV., shows that the concavity of the *a*-curve to the *y*-axis is not removed, though it is somewhat reduced, by the correction described. If the averages could be reliably carried further than to the fifty-sixth practice period—which they cannot because of some subjects 'falling out' with incomplete records—the *a*-curve would rise comparatively much more rapidly as the *e* values become very small; it would reach 400 when the actually recorded errors equalled zero. From this point the fraction, $\frac{1}{2}$, added to *e* should gradually decrease to zero, bringing the *a*-curve up to infinity. If fatigue were eliminated by the subjects' resting occasionally, it is possible that this point on the *a*-curve would be reached (theoretically) early enough to give the curve a decidedly concave aspect to the vertical axis. In the figure in question the *a*-curve agrees with this view by taking rather suddenly a marked upward trend. It would be interesting, with a large number of subjects practicing up to a high degree of perfection, to trace the trend of this curve out further, and to construct a rational mathematical curve to represent it as the 'normal curve' of average attainment in ball tossing for the value of $K = 200$.

Now since $K = ae + \frac{1}{2}a$, it is obvious that the constant practice in successive periods can be measured accurately neither in the number of catches alone nor in that of throws alone irrespective of catches. Since *e* begins with a large value and decreases gradually towards, but never reaches, zero, more practice is obtained in early than in later periods if a practice period is defined as 200 catches (as is done in this experiment). There is reason to believe, however, that the additional practice in early over later periods is small, if not actually negligible in amount.[1] Throwing a ball so

[1] It should be noted that less practice in the early practice periods than was given in the present experiment would have a tendency to make the average attainment curve rise more slowly and the error curve drop less abruptly, making the former curve therefore more concave to the vertical axis than it is. A good way to test the effect of the excess practice in early periods is to have a large number of subjects take practice periods of 200 tosses, irrespective of the catches, and then to arrange the results into successive groups of 200 *catches* each and compare the curve with those in Figs. 3 and 4, which represent averages by our method. Some procedure that is a mean between these two opposite extremes in method would doubtless be the best one.

that it cannot be caught likely gives but little practice in throwing one so that it can be and is caught. Moreover, the method here followed has been common in animal learning, and in all other experiments in which *something is to be done* once or more times for each practice. In all such experiments the constants, or 'equal amounts of practice,' set for successive periods are not strictly equal. Just how much this inequality affects results in any given case is not easy to determine. The difficulty arises from the fact that as practice advances errors and irrelevant and confounding movements are constantly on the decrease while the successful or correct movements are increasing in frequency and in rapidity of performance. Equal time periods of practice has been suggested and extensively used.[1] While this practice has many advantages it also shares the disadvantages here pointed out. A skilled juggler may toss and catch the balls two dozen or more times while the beginner fumbles about with possibly only one successful catch. Are such practice periods equal?

But the greater error comes not from the inequality of practice as such, but rather in the statements of results. This is a most deceiving thing. Total work performed, or average amount of work without (an arbitrarily defined) error, or with a given number of efforts, or within a fixed amount of time, becomes the unit and is represented usually by the ordinates. Abscissa units usually stand for successive practice periods however defined. Such definitions of ordi-

It is, of course, obvious that if our method errs in the direction of excess practice in early practice periods, that of equal tosses errs the other way in as much as wild tosses doubtless give less practice in throwing-and-catching than do tosses which result in catches. Our own method, as has been pointed out in the body of this paper, is more in harmony with the usual practice, and it also gives equal results—200 *catches*—in the several practice periods. My friend and colleague, Professor W. S. Miller, of the department of education, has kindly helped me to convert the data from one of his representative subjects by the equal-number-of-tosses method into average attainment for comparison with my own curves in Fig. 3. Only about the first half of the data from this subject were thus converted. So far as we went no noticeable difference from our own averaged results appeared. His practice periods, however, were 300 tosses instead of our own 200 catches. So the results are not strictly comparable. It would seem that the inequality of different practice periods by either method will show little effect in results so far as the mere practice differences are concerned. But see the next paragraph in the text.

[1] *E. g.*, Thorndike, *op. cit.*, 295 ff.

nate values presuppose a homogeneity of the data measured; but as has been hinted above no such homogeneity exists in the majority of cases. Practice in the early part of the experiment is vastly different from practice in the later part, or, stated differently, no account is taken of how the attainment measured is conditioned reciprocally by 'errors,' of whatever nature they may be. This difficulty is so great, it seems to me, as seriously to invalidate the conclusions arrived at as to the different absolute rates of learning in different stages of the learning process.

In ball-tossing one might define a practice period in terms of *throws*, say as 200, 300, or 400 throws, thereby avoiding the difficulty suggested above. But this does not really avoid it, for all throws are certainly not of equal practice-for-throwing-and-catching value. In the second place one is deceived by such a method when one comes to the definition of results. In one case, not yet published, of which the writer is aware 'the score,' where this method was followed, is the number of catches. Not only is it wrong to assume that a bad throw—such as the beginner is prone to make—gives as much opportunity for a catch as a good throw; but it is even more fallacious to count only the *absolute* number of catches as results, since the *ratio of catches* to the *possible number of catches* is constantly changing. Such results will always, it would seem, give, deceptively, a rapid initial rise in the curve and negative acceleration later. This is due to the fact that out of, say, 200 tosses only 200 catches are possible, and as the 200-catch limit is approached the proportion of catches to possible-catches approaches a limit of 1 : 1. The influence of errors (both of throwing and of catching, if they can be successfully distinguished) is therefore not only a variable quantity directly affecting the catching, but reciprocally and indirectly the number of errors determine the ratio of actual to possible catches. From the very nature of the conditions the limit (in terms of the score) is physically imposed and it must be approached asymptotically. We are bound under such conditions, because of this relation to the errors not being taken into consideration,

to get a curve of attainment which resembles the 'normal' learning curve! All of which is only another way of stating the fallacy of counting results, or the score, in absolute terms of attainment, or of errors, only.[1]

It should be noted that *error* as defined in the present experiment—a miscatch ending a series of catches—is practically the same as *trial.* It is perhaps more clear that results should be equal to the number of trials *times* the average amount of attainment per trial, than that they equal the 'errors' *times* the average amount of attainment. The reason that the term error is used here is that it is in common use in similar relations in psychological literature and that each trial ends by an error, or miscatch. It is also desirable not to confound trial with single tosses, or throws. Counting results in terms of single throws is liable to result in curves of a deceptive nature, as explained in the preceding paragraph, curves which seem, wrongly, to indicate that *learning* is at first comparatively rapid and that it later approaches a limit very slowly.

Because of certain widely current statements about the absolute rates of learning in its various stages, it seems worth while to carry this question of the significance of certain learning curves, raised in the ball-tossing experiment in particular, over into the more general field. Here prudence suggests, in view of the time that has been available for this study, that we proceed very tentatively. The aim is therefore rather to open up the problem and to stimulate further experimentation than to give it any appearance of finality. If anything like dogmatic statements appear, this is to be explained on the ground of brevity and of the belief

[1] The results of Mr. Batson, published in a recent *Psychological Monograph* which appeared after the present paper was sent to press, confirm the general position here taken. Mr. Batson has plotted his results in the three different ways discussed in this paper, and gets for error curves the rapid initial decline usual with such curves; for what we have called average attainment he gets curves which are concave to the vertical axis. Plotting a curve of the number of catches in successive practice periods with a constant number of *throws* he gets a curve which in general shows a rapid initial rise and a final asymptotic approach to the "physiological limit," or better to the 1 to 1 ratio. He does not, however, enquire into the meaning of the different kinds of curves.

that they may be more effective than highly guarded suggestions in bringing about this result.

An Examination of the Nature of Learning Curves

Why do our curves in this experiment, representing errors on the one hand and the average amount of accomplishment on the other, seem so contradictory? Both of them represent kinds of data (results) frequently used as bases of inferences as to relative rates of learning in the different stages of the process. Moreover, error curves and attainment curves are frequently compared directly without any kind of transformations. As has been pointed out in the specific cases considered, errors and average attainment have reciprocal relations to each other as dependent variables. Each factor conditions the other in such a manner that it is entirely unsafe to define learning in terms of absolute units of only one of these processes. In most cases, as was found true in ball-tossing, time curves and error curves have the same general form, so that the former have the same relations to the curves of average attainment that the latter have to them. In certain problems, such as maze learning, excess distance curves show the same general nature as time and error curves.[1] Of course, in these various cases 'errors' are not the *same kind of things*, even though they may be safely comparable. Much depends upon how they are defined, as is well known.[2] On the whole more errors means more time taken for a certain result in terms of performance; the same thing is true in a rough way of the relations of excess distance to performance or attainment. For the general purpose of this enquiry, then, we may confine our-

[1] Cf. Hicks, Vinnie C., 'Relative Values of the Different Curves of Learning,' *J. of Animal Beh.*, 1911, *1*, 138–156.

[2] This is true even within the same kind of experimentation. For example, in the case of animals learning the maze an error may be defined as merely putting the head into the *cul-de-sac*, as entering it with the full length of the body, or as going the full length of the *cul-de-sac*. A return in a single section without a bend may also be counted as an error along with any complete return, or the latter may be counted for as many errors as there were sections passed through in the return, or, a third possibility, no returns may be regarded as errors if the animal does not depart from the true path.

selves to error curves, keeping in mind that, with certain cautions, time or excess distance curves may be substituted for them without serious error.

The reciprocal relations of errors and average attainment, as we have defined these terms in ball tossing, are obvious. An error always ends a series. When the error count is large the average number of successive catches (the 'series') must be small, and conversely. Suppose that on the average only one catch can be made in a trial, or a series, at the beginning of the learning. Then the errors in a practice period of 200 catches must amount to 200. But an improvement or gain of one in each series (making an average of two successive catches per trial) will bring the errors down to 100, a change of 100 absolute units. An average of 3 catches each trial brings the errors to 66; of 4, to 50; of 5, to 40; and so on. If we plot these values on an absolute scale, representing errors by e and average number of catches by a, we get a 'representative learning curve' of the e-values when the a-values advance by equal absolute amounts. That is, the e-curve drops rapidly at first and then progressively less and less. This is of course on the assumption that we give these values the relations indicated in the equation

$$\frac{K}{e + \frac{1}{2}} = a,$$

and regard K as a constant. Very small changes in a will mean very large changes in e, and *vice versa*. Proportionately, of course, the changes are equal in both factors; *but learning curves are almost universally plotted in absolute amounts*, despite this fact. In the beginning of any learning process the values of *a must be small;* hence *e will always change rapidly on the initial part of the curve* thus plotted. When a gets rather large, e must change slowly, if the changes in a, absolutely, are constant. These changes in the error curve are obviously in no sense indications that the *learning* is at first rapid and that it then gradually decreases in rate.

Let us take a theoretical case of ball-tossing and assume that the increase in the average number of successive catches

—the *a*-values—is constant, adding one unit for each successive practice period. In such a case the *a*- and *e*-curves will have the forms shown in Fig. 5, when $K = 200$. When K is given a value of 40, the *e*-curve takes the form of e_1 in Fig. 6. The curves e_2 and e_3 correspond to gains in the

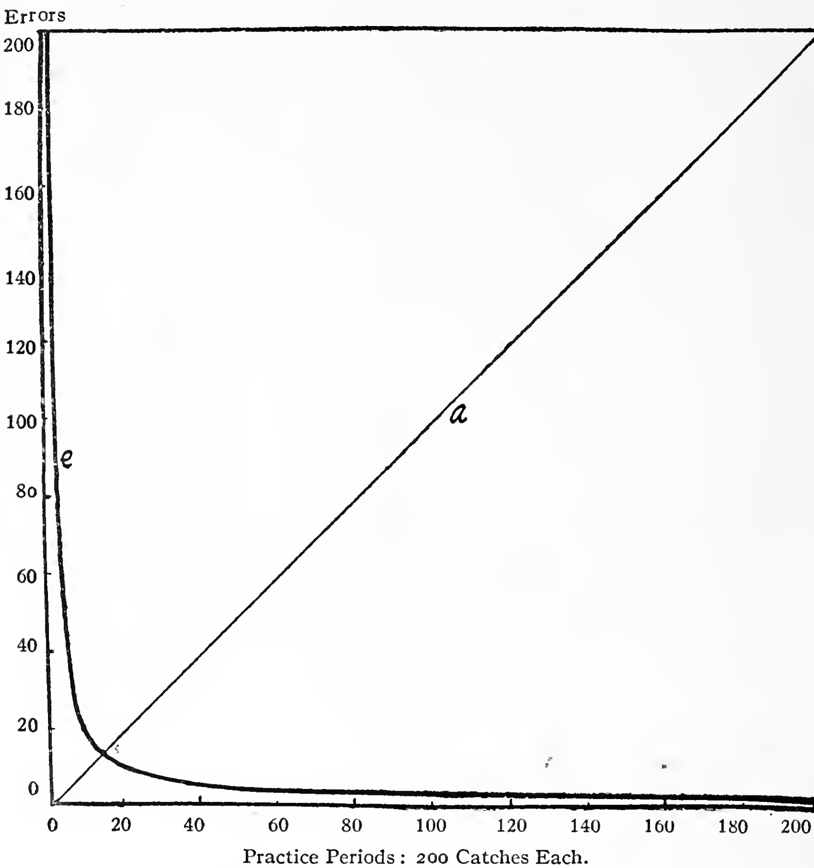


FIG. 5. Curve *e* is a theoretical error curve corresponding to an increase in *a*, average number of catches, of one catch in each succeeding period, when $K = 200$.

a-values of $\frac{1}{2}$ and $\frac{1}{4}$ units, respectively, in successive practice periods. These *e*-curves all have the general shape of the 'normal curve of motor learning,' falling off rapidly at first and more and more slowly as the learning reaches an advanced point. This type of curve is *normal* to motor learning

only because errors, and number of seconds (time), are the easiest data to get in such processes! This unfortunate circumstance, as it has proved to be, a mere contingency, seems to have led to a misapprehension of the real nature of learning. The relative rate of learning at its different stages has come to be judged, in the hands of not a few investigators, by the rate of decline of this curve. It should be noted that the smaller the values assigned to K the more gradual the decline of e. Relatively slow, constant gains of the a-values also affect in a similar way the e-curve. A rapid initial rise of a produces, of course, a more abrupt fall of the e-curve, as shown in Fig. 7, curve e.

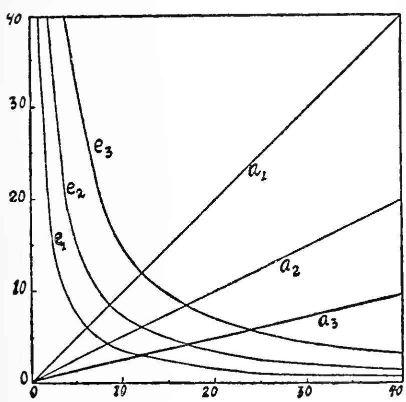


FIG. 6.

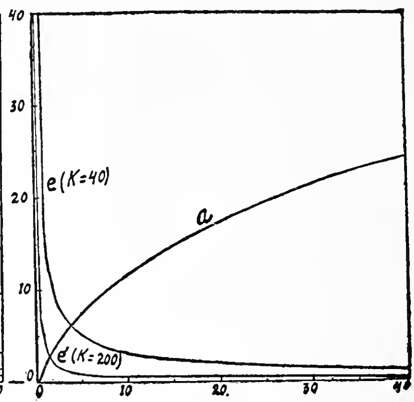


FIG. 7.

FIG. 6. e_1, e_2, and e_3 are theoretical error curves corresponding respectively to equal absolute successive gains of one, one half, and one fourth in the average numbers of catches, a_1, a_2, and a_3. $K = 40$.

FIG. 7. e and e' are theoretical error curves derived from the average attainment curve a, when $K = 40$ and 200 respectively.

The *form* of the a-curve is a third factor that affects the reciprocal e-curve. If, for instance, a begins with a value of 1 and increases in each succeeding period by 1/10 of its value in the next preceding period, we get curves like those shown in Fig. 8, when $K = 200$. These curves resemble very closely those actually obtained in ball-tossing by certain subjects. Whether this type of a-curve, representing *equal proportional gains*, comes nearer being a constant progress in

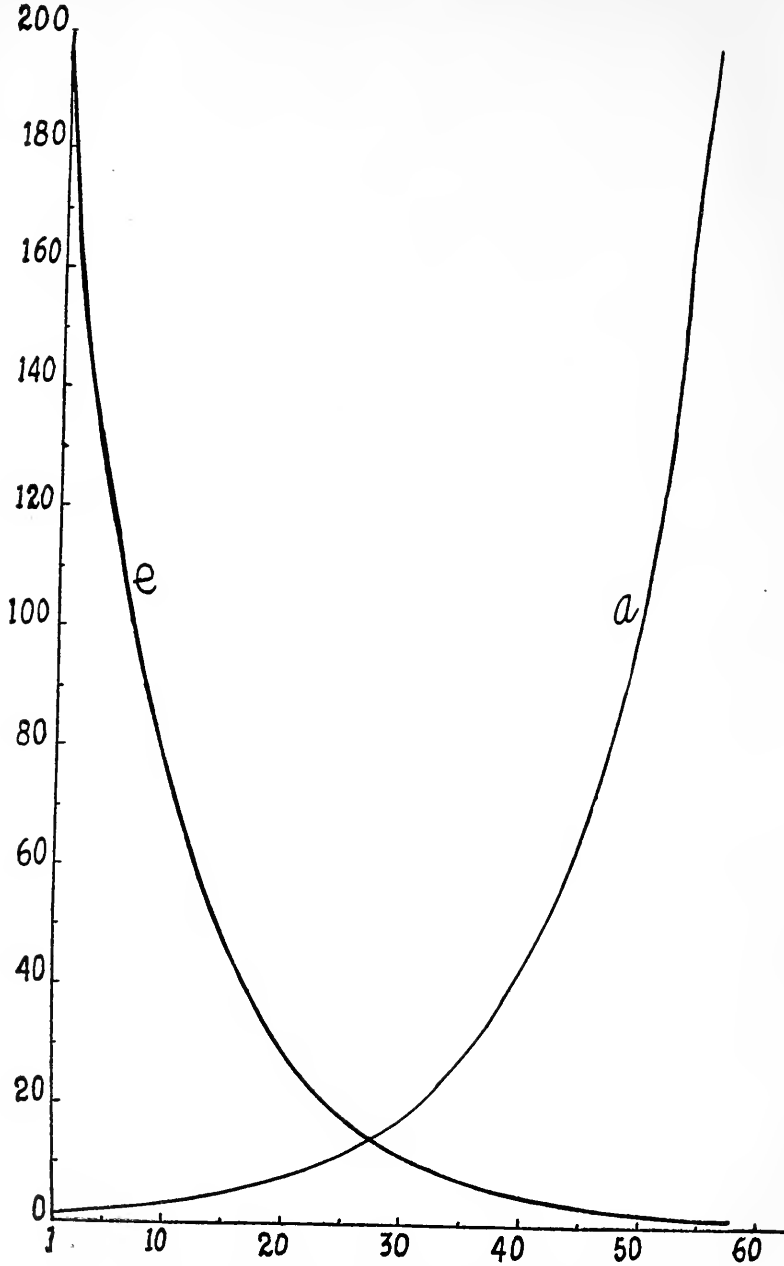


FIG. 8. Curve *a* represents equal proportionate gains of 1/10 over each preceding period, the initial value being 1. *e* is the corresponding error curve when $K = 200$.

learning than equal absolute gains of the a-values is a question that may be answered differently by different authorities and by the same authority for different learning processes. At any rate when plotted in absolute units the curve of such a gain is concave to the vertical axis, as are the average attainment curves of ball-tossing. This correspondence might suggest that learning in ball-tossing advances by approximately equal proportionate gains; but to take this stand would be to get into a questionable position, to make the mistake with a slightly different application which has been made to date by certain psychologists in the interpretation of error and time curves. When the average attainment is great the errors are few, and because of the reciprocal relations of a and e, pointed out in the foregoing, *slight changes in errors must mean large changes in average attainment*, when these values are plotted in absolute units.

If the a-curve has an initial rapid increase, as shown in Fig. 7, making a convexity toward the vertical axis, the initial drop of the e-curve is increased in amount. If the K in our equation, by which the e-values were derived, is decreased in value the initial drop will decrease correspondingly. (Compare curves e and e', Fig. 7.) It is of course obvious that any tendency to a rapid initial increase in average attainment, as defined in this paper, will correspondingly increase the decline of errors, or trials.

When we consider these relations of curves of different kinds of data obtained in learning it becomes evident that great care in the comparisons of different learning curves is advisable. All 'errors' and 'scores' are more or less arbitrarily defined, and it seems necessary for comparative purposes that some standardized methods of conversion of data be devised. At present we are forced to recognize the insecurity of any conclusion as to the absolute rates of different stages of learning, based on such curves of learning as are available. As a rule curves of learning are based on the total amount of accomplishments of the kind specified as 'results,' but in many cases they are constructed on the basis of what *can be done* in a stated number of minutes

after certain multiple units (weeks, hours, or minutes) of practice. In such cases the initial stages do not allow of as much practice *in the particular aspect of learning that is plotted* as do later stages, for more time absolutely is lost on various conditioning functions. The considerations of some of these matters may even result in the view that a relatively more rapid initial learning takes place in certain cases than is commonly assumed from the nature of the curves of the data obtained. On the whole, however, the evidence for the general statements current, that learning is initially rapid and gradually decreases in rate to a physiological limit, is far from satisfactory. Meanwhile it must be recognized that the practice of putting into graphic form the 'progress' in different learning processes studied, has had fruitful results toward a more quantitive psychology of learning and toward the establishment of a better knowledge and control of various aspects of learning in educational fields. It may be asked, May we not still continue to plot curves of various learning processes and to say certain definite things about them? Yes; provided we are careful as to *what* we say, and as to what comparisons we make. The curves are not the learning process and in most cases they represent it but poorly. For example, plateaus may not be plateaus *in learning* at all; we may be plotting some aspect of the process which for the time is neglected for some other aspect. On the other hand, a real plateau on the latter part of an error curve, where change is slow, may be practically obscured, while it would have been much more conspicuous in another type of curve.

Some Applications and Suggestions

In every case of learning scientifically studied some definite units of practice must be decided upon—food must be obtained once or several times by an animal either escaping from or getting into a problem box, or by one finding its way past a certain number of blind alleys to the food box in a maze; a given number of catches of balls, of additions of numbers, of throws of javelins, of substitution of symbols,

or of distribution of cards must be made; or, a more comparable and constant procedure in many cases, a given number of minutes or hours must be devoted to a certain specified kind of practice. Whatever the unit of practice may be we may designate it, with certain cautions, by a constant, K.

The score, or results, at any stage of the learning may be determined in at least two, and probably only two, general ways. First, the number of errors (or trials)[1] made in performing the 'work' designated by the constant—such as errors in getting to the food, in making 200 catches of the balls, in striking within certain limits in javelin throwing, in adding certain groups of numbers, in tracing a star, etc.—may be recorded; or the number of seconds required for some such performance may be indicated;[2] or we may in the maze problem take the excess distance covered. In the second place, we may choose the reciprocal results, the amount of performance in certain units (determined by the nature of K) which can be reached on an average without an error or in a given number of seconds. The first of these classes of data on results will give what we have called e-curves; the second, what we have in a general way spoken of as average attainment curves. In the following very tentative comparison of curves taken almost at random from various important experiments on learning by different investigators we shall try to keep distinctly in mind these different factors. In doing this we in no way claim for our classification and formula superiority over any others that may be proposed. We shall use K always to designate the constant of practice; e, the errors, number of seconds (time), or excess distance in maze; a, the amount of performance, attainment, or average attainment. For the transformations of the data from various experiments we shall take some form of the

[1] See *supra*, p. 205.

[2] In experiments on typewriting, telegraphy, crossing out letters or numbers, etc., an error may be defined as failure to make two successive strokes within a given fraction of a section. This would not be greatly unlike our errors in ball-tossing, since too much attention to the case at hand would make an error in the next one. Here again time and error curves show close relationships.

equation,

$$\frac{K}{a} = e,$$

to suit the particular conditions of the case in question. Our purpose is to attempt schematically, under some one of the two classes of results defined, a comparison of data from different experiments. It is obvious that different values of K, whether stated numerically or otherwise, and different arbitrary definitions of e, particularly when e designates errors, will wholly determine the values of a and therefore the fruitfulness of our comparisons.[1] Our purposes are, however, chiefly to illustrate a procedure that is possible if the various experiments are originally planned for comparative purposes.

In most reports of experiments only two of our factors, K and either e or a, are given, and from these the other one—even if only a theoretical factor—may easily be derived by the simple arithmetical process indicated by our equation. We may thus in a general way compare different 'kinds' of learning and see whether there seems to be so much difference as has generally been supposed, or we may at least suggest a method for comparisons in future experimental work. In our illustrative cases considerable difficulty has been experienced because of the fact that data were not fully given in all cases. To facilitate comparisons we will make all e-curves represent decreasing values and a-curves increasing values. This will require in some cases simple transformations of some of the authors' results before the application of our equation is made, but it will not change their values. In every case the curve represented by a continuous line (———) is the curve giving the original data of the author, and the curve in a discontinuous or broken line (– – – – –) is the one supplied by the present writer through dividing the original values (the a's or the e's) into the constant, K. The broken

[1] That is, even for purposes of comparisons between different experiments of the rates of change in e- or in a-curves, we must, to make our conclusions safe, keep the other two factors of the equation, K and a (or e) approximately equal in both cases. We must also be sure that our e (or a) values compared are comparable.

line therefore represents only another statement of the results supplied by the original author, one, however, which may be more directly comparable with results of other experiments.

Since it is manifestly impossible to rearrange the data from most of the experiments referred to so as to get accurately comparable values of K, a, and e, we can expect only a general conformity of the curves compared. More quantitative comparisons must come from results of experiments deliberately planned for comparative purposes, thus giving the different factors of the transformation equation as nearly equal values as possible. The relative size of the abscissa and the ordinate units is another matter that greatly determines the form of the curve, as is well known; but with unequal values of the factors of our equation it is manifestly incorrect to make these units in the several cases all equal. This makes it clear that our comparisons must be regarded only as illustrative.

In Fig. 9 curve *a* represents the average number of catches without a miss in practice periods of 200 catches each, by Swift's subject *A*, and *e* is the error curve, corresponding to the error curves of our own subjects. The data for the average number of catches was taken from Swift's curve,[1] and it was rearranged into such a form that successive practice periods are equal so far as the number of catches are concerned. As is well known the actual practice periods in Swift's experiments were far from equal. The error curve, whose values were obtained by dividing the successive average number of catches into 200, corresponds very closely to many of our own error curves. The learning of Swift's subjects has usually been considered to differ from the normal type of learning. There is evidently little ground for this view when similar results are compared.

Bryan and Harter's classical receiving curve for learning telegraphy is represented by *a* in Fig. 10.[2] A corresponding time-curve is represented by *e*, the number of seconds required for 'receiving' five letters. In this case a difficulty

[1] *Amer. J. of Psychol.*, 1903, *14*, 201–251.

[2] *Ibid.*, 1899, *6*, p. 350. Subject W. J. R.

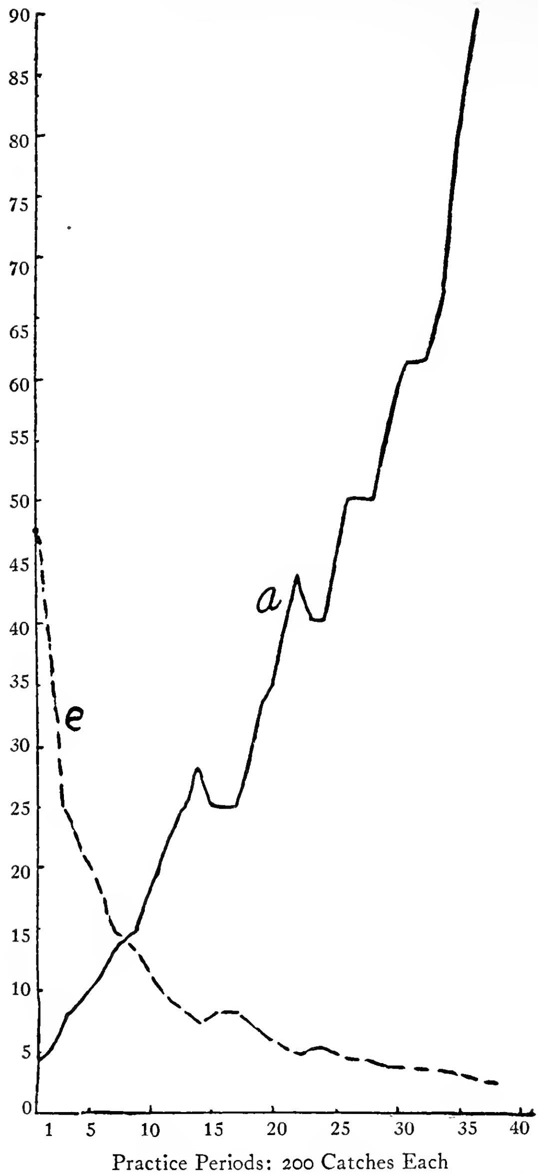


Fig. 9. Error curve, *e*, and average-number-of-catches curve, *a*, of Swift's subject *A*, rearranged into practice periods of 200 catches each.

is encountered. The original curve, *a*, of the authors represents the number of letters per minute *after the several successive weeks' practice*, and is not an average of the whole practice period. It seems probable that with a proper adjustment of ordinate and abscissa units, and with a slightly different tabulation of data—which is impossible for us to

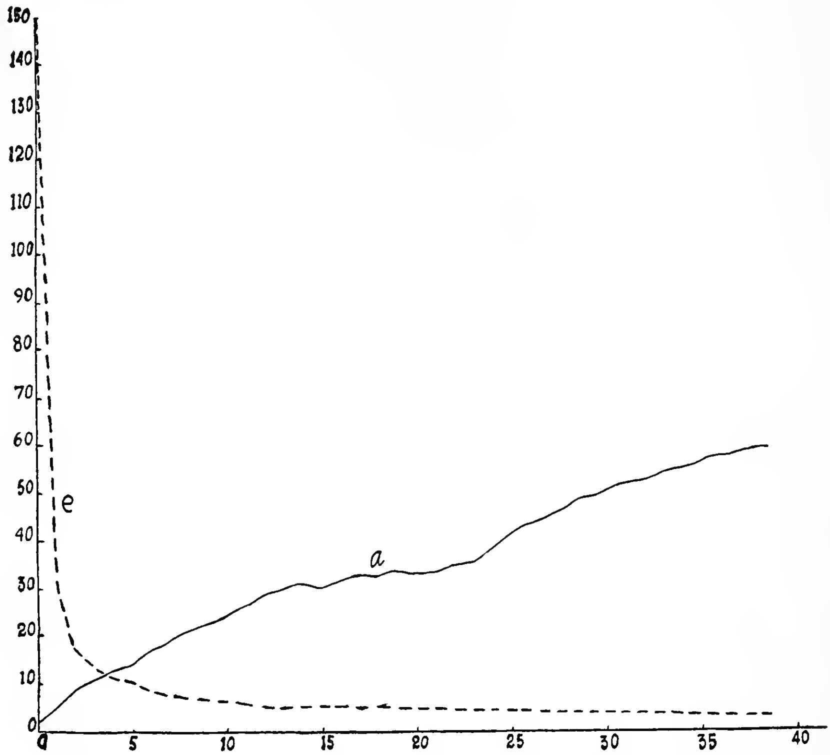


FIG. 10. *a* is Bryan and Harter's receiving curve (number of letters per minute), telegraphy experiment, subject W. J. R. *e* is the corresponding time-curve, representing number of seconds for receiving five letters. Abscissa unit = one week of practice.

show because the complete data are neither given in the original report nor embodied in the curve—we should get the 'normal (time) curve.' The curve here given from such data as was available has an initial drop that is very sudden. Evidently, however, this is the curve that we must compare with other time and error curves, not that (*a* in the figure)

originally given by the authors. It is interesting to note that the plateau, so marked in the original, or average, curve, is not noticeable in the time curve, because it occurs where absolute changes are very small in this curve, on the principle already explained in this report. This may be a valuable suggestion as to the reason why some experiments on learning do not show, and others do show, plateaus. Speculation here without further data, however, is unprofitable.

In animal psychology it has become customary to regard 'motor learning' and 'sensory discrimination learning' as somehow different in *kind*. As is well known, both kinds of learning are frequently represented in terms of errors. But their error curves are markedly different. To quote from Yerkes: "The former [error curves of the labyrinth tests] fall very abruptly at first, then with decreasing rapidity, to the base line; the latter [those for the discrimination test], on the contrary, fall gradually throughout their course. Evidently the labyrinth habit is more readily acquired by the dancer than is the visual discrimination habit. Certain motor tendencies can be established quickly, it would seem, whereas others, and especially those which depend for their guidance upon visual stimuli, are acquired with extreme slowness."[1] Watson, in 'Behavior,' seems to take the same position in general, making different divisions for discussion of sensory and of motor habits.[2] While it is, of course, not to be denied that sensory discriminations may be made very difficult for an animal, so that learning to discriminate will be slow comparatively, it should not be overlooked that the same thing is true of 'motor learning' and that a difference based on sensory and motor habits cannot properly be made. The distinction has always seemed to the writer to be artificial and based on bad psychology. Any habit may be regarded as either sensory or motor, or better, both, according to the point of emphasis. This was long ago made clear by Dewey in his well known article on the reflex arc concept.

One of Yerkes's sensory discrimination curves,[3] modified

[1] 'The Dancing Mouse,' p. 236.

[2] Chapter 6.

[3] *Op. cit.*, p. 230.

to represent increase over 50 per cent. correct responses, as explained in the foregoing, instead of decrease from 50 per cent. error as stated by the author originally, is shown in Fig. 11. Here it is the usual average attainment curve, as might be expected from the data which it represents, even though it is called an error curve.[1] Now, on the basis of

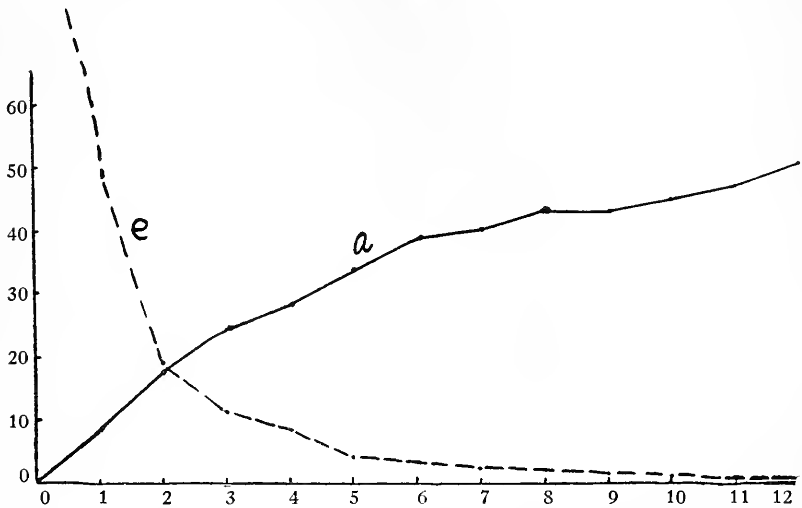


Fig. 11. Curve *a* is rearranged from Yerkes's curve of brightness discrimination by the dancing mouse. *a* represents per cent. correct choices above 50 out of ten trials. *e* is the corresponding (theoretical) error curve derived with a value of $K = 50$.

our equation it is possible to get a value corresponding more nearly to the usual error curve of so-called motor learning. For this purpose we have used the equation

$$\frac{K}{a} - 1 = e,$$

so that the error curve when $a = K$ will reach the zero line, as would actually be the case. Taking $K = 50$, the highest per cent. correct above 50 that is possible, and getting our values of a by subtracting the values given in the last column

[1] This statement, it is frankly confessed, needs more study and analysis. It is not easy to see in the present case the values comparable to our *a*- and *e*-values; but the curve given by Yerkes is certainly not an *e*-curve, as here defined, and should not be directly compared with the error curve in the maze learning as is done. If Yerkes had compared it with some sort of attainment curves he could not have come to the conclusion that he expresses.

of Yerkes's table from 50, we obtain the *e*-curve shown in Fig. 11. Whatever this *e*-curve represents concretely in the present discrimination experiment, it is certainly interesting to note that it is almost identical with Yerkes's own error curve of the labyrinth test.[1] Of course, this may not be wholly convincing since the ordinate and the abscissa units are not respectively equal in the two cases, but it is very suggestive of a better mode of comparison of curves than has hitherto been used, one that should avoid some of the errors made in comparisons in the past. It is, of course, always to be admitted that in spite of any manipulation of data one learning process may be made very much more difficult than another. By the use in comparative psychology of some such conversion method as here suggested, however, we may avoid the errors of making misleading comparisons between curves which may chance to be called by the same name and thus conclude that there are different kinds of learning, and we may also avoid erroneous statements regarding the absolute rate of learning at different stages of the process.

Murphy, in a recent article,[2] shows the usual confusion between error (or time) curves and average attainment curves. His subjects practiced at throwing the javelin. Five (?) throws were made by each subject at each practice period, these periods being distributed differently for different groups. Results for each practice period were given in terms of the average distance in centimeters from the bull's eye. "In the above experiment," the author concludes, "the curves did not follow the normal curve of learning to any marked degree; but those in the alternate day group generally gave a better approximation to a regular learning curve."[3] As a matter of fact, Murphy's curves fluctuate a good deal because of the comparatively few trials in each

[1] 'The Dancing Mouse,' p. 235.

[2] Murphy, H. H., 'Distribution of Practice Periods in Learning,' *J. of Educ. Psychol.*, 1916, 7, 150–162.

[3] *Ibid.*, p. 161. Murphy plainly refers to the error, or time, curve as the 'normal curve of learning.' Such curves will be found on examination almost universally, if not quite so (even in spite of the fact that errors are differently defined) to correspond rather closely to our *e*-curve in the foregoing. It is rather surprising that this fact has not received explicit attention before, as it does not seem to have done.

practice period. Some of them, however, show a fair approximation to the average attainment curves usually found. His own results were plainly average attainment. A bunching of his results into more practice for each period—a process impossible to us because he does not give individual results—would obviously give more uniformly declining curves. If, then, the results were stated for successive periods in terms of differences from the largest average distance from the bull's eye (*i. e.*, from an error of 82 cm.), we should have a gradually rising average attainment curve very similar to the usual *a*-curve as defined in this paper. It is unfortunate that Murphy's results are not given in enough detail so that the transformation suggested here could be made as illustrative. To compare curves of average attainment, as Murphy has done, with the 'normal curve' (of error or time) is to make an unfortunate error that is all too common in psychology and education. All curves of different aspects of learning have frequently been regarded as somehow representing the same thing—*the* learning process—however different the data may be which they represent in the various kinds of practice.

In the javelin-throwing experiment it is not at once obvious what corresponds to errors in an experiment such as ball-tossing or the learning of the maze. Let us suppose, however, that every throw missing by more than a given distance, say 10 cm., had been called an error, and that every hit within this limit had been counted a score. Under these conditions if each practice period consisted of some appropriate number of hits (hits within the arbitrary limit) it would seem probable that the results would be comparable in a satisfactory way with those of our ball tossing experiment, for instance. It is true that the correspondence of the data from two such experiments would yet be only approximate in as much as the balls tossed and not caught would give relatively less practice in tossing-and-catching than throwing a javelin slightly outside of the 10 cm. limit would give in throwing-the-javelin-within-the-limit. It must be noted that in our own experiment the successive practice periods are equal in amount of practice only in as far as

throws that are caught are counted *as practice.* Though there is some degree of justification in not counting as practice balls thrown wildly and missed, it is true that many balls are missed through the merest slip in technique, such as letting one ball strike the edge of another when a fraction of a millimeter to the side would have resulted in a catch. Certainly there is practice in such a throw and miscatch and one learns by it, among other things, to avoid such collisions.

In comparing results of different kinds of learning we must exercise considerable care not only to get curves of the same general type, as above illustrated, but also to see that practice is approximately equal in successive periods in each of the experiments compared and that the practice periods of each kind of learning are about equal. The writer hopes to find opportunity in the near future to obtain data of more nearly comparable values in two or more kinds of experiments than exist at present. It must be emphasized, however, that errors resulting from slight inequalities of practice periods are really less serious than those from errors in arrangement of data, some of which have been briefly and very imperfectly pointed out in this paper. It is unfortunate that important comparisons of incomparable data, and that classifications of different 'kinds' of learning have worked into psychological literature and thought with scarcely any attention to this matter of basic import to any reliable comparisons among learning processes in different kinds of performances.

A single result of taking time, or error, curves to signify changes in the learning process itself comparable to changes in the curves may be found in Hobhouse's early reaction[1] to certain statements of Thorndike. Thorndike showed time curves of various animal-learning processes and contended that learning resulted from gradual physiological changes in the nerve tracts. Hobhouse, it will be remembered, objected that the curves in question were in reality steeple-like, the sudden drops being indicative of the influence of something like ideas, or 'practical judgment.' Thorndike did not apparently make the mistake of assuming a corre-

[1] *Cf.* 'Mind in Evolution,' 1901, 145 ff.

spondence between the curves and the actual learning, and his general position was against such an assumption; but he did not explicitly show why the rapid initial drop occurred. He has more recently in the volume already referred to, for example, given various reasons why learning should be more rapid initially than later in the process. These reasons may, indeed, apply in certain cases; but certainly the whole matter needs further experimental investigation. Hobhouse's arguments, so far as they were based on the steeple-like features of the curves, were certainly groundless, though critics of his position have been slow to point out this fact, apparently accepting his own assumption. Numerous statements might be quoted, besides those already given, to show that this has been the usual understanding of the significance of learning curves.

Summary of Results and Conclusions

As a result of experiments on ball-tossing with twenty-eight subjects it has been found that the learning curves of ball-tossing are not different from learning curves of other processes, provided that similar kinds of data be compared. The error curve of ball-tossing, for example, has a rapid initial decline as have error curves—the usual 'normal curve'—of various other types of learning. The curve of the average number of catches per trial, on the other hand, rises gradually at first, making almost a straight line, and more rapidly later. This gives it a concave aspect to the *y*-axis, as has been noted by Swift. Neither of these curves, however, is to be taken as exactly indicative of the relative rates of learning at different stages of the process. Other forms of plotting the results are possible, some of which would seem to suggest that learning is initially rapid and becomes slower as the process advances. Serious errors have resulted in the interpretation of learning from giving too much significance to the forms of any of these curves. On the whole little evidence exists, so far as ball-tossing is concerned, for the current view that learning is at first much more rapid than later when considerable skill has been attained, this despite the form of the error curves. Error curves are

found to be mathematically so related to curves of average attainment that the former must drop comparatively very rapidly at first while the latter change slowly, and, conversely, they must drop very slowly later while the average attainment curves rise rapidly. This is in ball-tossing. The changes described do not mean corresponding changes in *the learning.* There is good ground for the view that the same thing may be true of all learning curves. In general, it appears, curves of errors and of time (and of excess distance in the maze) will, because of this mathematical relation to attainment curves, always have a sudden initial drop. In cases showing a rapid initial rise of the average attainment curves it is probable that learning is initially more rapid, as has commonly been supposed, but the difference in the initial and the later rates is likely not as great as it has been thought to be. Even in such cases, and ball-tossing does not seem to be one of them, the evidence is questionable—the rapid initial progress may in many cases not be one of learning at all, but one resulting from the method of stating the results. The whole subject needs a careful experimental overhauling and methods of making data from different experiments more comparable need to be worked out. A tentative procedure is suggested in the present paper, and a few applications are given as illustrative, but only roughly, due to the incompleteness of reported data by various authors.

The basis on which our present views of plateaus and of the approaches to physiological zero rest is insecure. It is found that whether or not a plateau appears and the degree of its occurrence, depend much on the type of curve chosen to represent it. If, for instance, it occurs late in the learning process it may be almost wholly obscured in the error or the time curve. This may be one reason why error and time curves so seldom show plateaus. When the proper conversion of data is made before comparisons there seems to be little ground for the distinction frequently made between sensory discrimination learning and motor learning in animal behavior. The difference is likely mainly one of the method of calling out the reaction and of stating the score; the learning in both cases can be so stated as to make the two processes quite comparable.

THE SPEED AND ACCURACY OF MOTOR ADJUSTMENTS[1]

BY JOHN J. B. MORGAN, Ph.D.

Princeton University

Motor activity must necessarily consist of habitual responses to the same or repeated stimuli, unusual responses to novel stimuli, or modifications of habitual responses to a stimulus partly old and partly new. As it is doubtful whether an entirely novel situation ever presents itself to an adult human being, and as stimuli are seldom repeated in exactly the same combinations or successions, our activity consists for the most part of modifications of our responses to meet changes in the stimulus complex.

The more nearly identical the complex of stimuli, the more nearly identical will be the responses; as stimuli become more diverse, the responses must be more radically adjusted. For instance, when one walks along the street the locomotor stimuli are practically the same. If the steps were accurately measured, however, they would be found to vary in speed, in length, in direction, in the intensity of the foot-fall, in the height to which the feet are raised, and probably in other ways. These minor variations are due in all probability to slight changes in the situation complex such as irregularities in the sidewalk, changes in the position of other parts of the body, breathing, blood flow, glandular activity, etc. A greater variation occurs, or what is the same thing a more radical adjustment is made if one steps on a sheet of ice, comes to a gutter, or is suddenly confronted by another pedestrian.

Historical

The three directions in which the adjustment of a movement may take place, in the physical sense, are extent, time and force; and these three must work together, a change in

[1] Contribution from the psychological laboratory of Columbia University.

one being compensated for by a change in one or both of the others. These three functions of movement have been adopted by psychologists in their study of bodily movements. The problems attacked have been whether we may perceive time, extent and force independently, or whether one or more are fundamental and the others derivative; and the relative accuracy with which the different factors may be perceived and controlled. (13, Chapter III.)

The experimental work on these problems has been of two sorts. In some of the work the attempt was made to eliminate one factor and study the relations of the other two, while some investigators have let the subject make free movements and have secured measurements of all three. Most of the work has been done in connection with the study of lifted weights, while studies that have had to do primarily with movement have been mainly concerned with the analysis of constant errors.

Müller and his pupils (8, 10, and 11) were primarily concerned with the judgment of lifted weights. They controlled the extent to which the weights were lifted and measured the speed with which the different weights rose and by this method endeavored to analyze the factors essential to the judgment of lightness or heaviness. Besides their findings that have to do peculiarly with lifted weights, they brought out one fact that seems to have a wider application; that is, the fact of a motor set (Einstellung). It is a fact that if a heavy weight is raised a number of times and then a light one lifted, the light one will be lifted much more quickly than were the heavy ones. This fact led them to emphasize time as a factor in the judgment of lifted weights. They showed that if two objects of equal size but different mass, such as 700 grams and 2,500 grams, were lifted alternately with equal speed and to the same height for about thirty times, a set is obtained such that, if for the 2,500 gram weight another is substituted of 850 to 950 grams, it is judged lighter than the 700 gram one. The explanation given was that the effect of repeatedly lifting a light weight followed by a much heavier one is to establish a set of the organism such that greater force will alternately

succeed a less amount. If, after a habit of alternating the amount of force has been set up, a lighter weight is then substituted for the very heavy one, it will be raised with such force that it will rise more quickly than its lighter companion and be judged the lighter.

Other evidence that speed is a fundamental factor in the adjustment of motor response has been brought forward by Jacobi (5). He had his subjects raise simultaneously two weights, one with either hand, and judge which was heavier; the moment each weight rose being recorded graphically on a smoked drum. He wanted to see whether the inertia of the weights was not a basis of comparison. He found that when the weights were judged the same they began to rise at about the same time, but when one was judged heavier it took as a rule longer before it began to rise. When the difference in time was less than 0.08 sec. the two weights compared usually seemed equal, between 0.08 and 0.12 sec. they sometimes seemed equal, but when the time was greater than 0.12 sec. the weight raised last was always judged heavier, whether it was in reality heavier or not. He concluded that the comparison of weights consisted in the comparison of the times necessary to overcome the inertia. Woodworth (13, p. 135) objects that his results were too few to serve as a basis for any statistical correlation and hence lacking in conclusiveness.

Experiments with the size weight illusion also show that inertia as well as the height to which the weights are lifted feature in the judgment. Claparede (2), experimenting with this illusion, used three cubes 8, 12, and 16 cm. on a side (that is, with volumes of 512, 1,728 and 4,096 cubic centimeters), each weighing 345 grams. Each weight was surmounted by a ring with which it was raised, and by means of which an electric contact was made which indicated when traction began. This with the record when the weight rose represented the time it took to overcome the resistance of the weight without lifting it. He also obtained a record of the height to which the weights were raised. He found that on the average it required 0.12 sec. to overcome the inertia of the largest weight, 0.21 sec. for the medium weight, and 0.62 sec. for the smallest

weight. The largest one was raised 25 mm., the medium one 20 mm., and the smallest one 10 mm. As the illusion causes one to judge the larger weight lighter we have a correlation between the speed with which the inertia is overcome and the judgment of lightness or heaviness, and also a correlation between the height to which they are raised and the judgment; the weight that is judged light being lifted more quickly and to a greater height than the one judged heavy.

These experiments seem to show that before a movement is begun the individual establishes a set to correspond to the anticipated resistance. The anticipation may be of the nature of a habit set up by some such procedure as that of Miss Steffens (11), or it may be by the comparison of the size of the weight to be lifted with the one previously lifted. A change in the speed of the second lift, in the distance to which it is raised, or in the latent time required to overcome the inertia gives the basis for the judgment of lightness or heaviness.

When the compared weights differ greatly it is likely that another factor comes into play, namely a change in the amount of musculature used in the two cases. Révault d'Allonnes (1) obtained graphic records of the struggle which goes on when one tries to lift a weight that is heavier than the lifter anticipates. He arranged an apparatus by means of which a weight is held on a board suspended from a spring so that the slightest change in the force exerted could be graphically recorded. Besides this method he used a false weight which consisted of a jar with no bottom fastened through a table to a very stiff spring, so that it could be raised only with great effort. By the side of this false weight he placed similar appearing weights of differing mass, ranging from 670 to 30,550 grams. The subject was told to estimate the weights in the order in which they were placed, which was from the heaviest to the lightest, and last of all the false one. The graphic records show that in the attempts to raise the false one several trials were made, each one stronger than the preceding. The time of this series of efforts ranged with different subjects from 6 to 18 seconds with an average of 10.05 seconds.

Fullerton and Cattell (3) oppose the Müller-Schumann theory, that time is the basis of judgment of lifted weights, on the ground that their experiments showed that force is more accurately judged than time, and that the perception of force more nearly follows Weber's law than does the perception of time. Besides they found that a person can arbitrarily vary the speed with which the weights are raised and yet judge correctly. Their experiments on the perception of force were complicated with extent so that their first criticism is not unanswerable. They used a spring dynamometer and in all except the heaviest pull the force was a direct function of extent. As to the second objection, the fact that one can arbitrarily change the speed and still judge accurately does not show which is judged more accurately, the force or the time, it only shows that judgment is not confined to time estimations.

Woodworth (13) investigated the relation of extent and force by means of blows. He arranged a lever so that it was operated by a string passing through a pulley to the hand. When the hand descended toward the table it raised the lever and when the hand struck the table the lever continued to move, its extent of movement after the blow being proportional to the force of the blow. He eliminated extent by two methods: one was to place a rod above the table to act as a starting guide to the downward stroke of the hand, the other was to allow the subject to see by the graphic record how far to raise the hand. He however had no control of the time factor and therefore his work shows nothing as to the relation of time and force. He found that gross changes in the force of a blow were accompanied by changes in the extent, but there was no correlation between the smaller changes in force and extent. He gives the following records of the proportion of cases in different experiments in which the finer gradations in force were accompanied by corresponding changes in extent: 62.8 per cent. of 670 cases; 54.6 per cent. of 928 cases; 52.6 per cent. of 207 cases; and 61.7 per cent. of 2,115 cases. The control of the extent by either of the methods that he used did not on the whole increase the accuracy of the force of the blows. He concludes that force is not closely dependent on extent.

We have, in short, one group of experiments which tend to show that our judgment of weights is not based on the amount of force required to lift them, that force is determined by the set of the musculature, but that the judgment is based upon the speed with which they rise or the distance to which they are raised. On the other hand we have experiments that would show that force is an independent entity which may be judged and controlled independently of time or extent.

Experimental

Previous experiments by the writer (9) have shown that, when a subject is instructed to use maximum force throughout an experiment in which various weights are lifted by means of a rope, a change in the mass of the weight to be lifted is not compensated for by the speed of the pull (extent being constant) so as to keep the resultant physical force the same. The results of these experiments showed that the instruction to use maximum force resulted in the subject making the pulls with different weights in nearly the same time, which of course means a marked change in force when the weights are changed.

The first part of the experiment to be described in this paper deals with the *speed* of this adjustment to a change in load, and the second with the *ability* to compensate for changes in load by proper changes in time, so as to keep the resultant physical force the same. The first part was done with the instructions to pull at one's maximum throughout, and the weights changed without the subject's knowledge of the time of the change or of the direction or amount of the change. In the second part the subject was told of the change and allowed to test the weight before making the pulls.

The question which led to the first part of the investigation (that of the speed of adjustment) was whether the adjustment was deliberative, reflex, or a local muscular phenomenon. It was thought that the determination of the speed with which the adjustment was made would throw some light upon this question.

To solve this problem it was necessary to have some mechan-

ism which would permit the subject to raise the weights with great speed and which would give a graphic record of the entire movement. To a carriage which ran in grooves hollowed in horizontal guides was attached a 100 vs. fork in such a position that it would write during the movement of the carriage on a flat smoked surface placed under it. From one end of the carriage a rope was passed through the partition into the next room and attached to the handle which the subject pulled. From the other end of the carriage a rope passed over a pulley on the end of the table on which the guides were fastened, then down and through a movable pulley to which weights could be attached and then up again to a fixed support. By this arrangement the weight moved just one half the distance that the recording carriage moved, and in computing the scores the fact had to be taken into consideration that the acceleration of the weight was one half that of the actual graphic records. As the force equals the product of the mass and the acceleration, what amounts to the same thing is to calculate by the usual formulas using only half the weights in the computation. The weights used, including the weight of the pulley, were 15,540 and 4,880 grams and in the calculations were considered as 7,770 and 2,440 grams.

Eight subjects were used, five of whom knew nothing of the nature of the problem to be studied, while three did. None, however, could know when the weights were to be changed as they were in a separate room, nor could they know the values of the weights. The interval between pulls was the same throughout the experiment, thirty seconds in each case, so that no clue could be obtained from the length of the interval as to the changing of the weights. The subjects were instructed to keep the rope loose between pulls, but as an extra precaution against any index that might be given by the tension of the rope it was knotted where it passed through the wall so that the tension would appear the same. Besides the carriage was locked between pulls till the moment the signal was given to pull, hence the subject could make no tentative pulls before the main one. The signal was given by means of an electric buzzer.

Each subject was given the following directions: "Your task will be to grasp the handle and pull each time you hear the buzzer sound. Pull as far as the rope will permit, exerting all your force throughout the movement. The time taken to respond to the signal will not be measured, so be sure you are perfectly set for the pull before you start it. The force you use is the thing that will be measured, so do your best every single time. Between pulls let the rope hang loose and start each pull from a position which permits the rope to sag."

The first five pulls were allowed as practice, were with the light weight, and were not used in the tabulations. After the first five pulls they were given in the following order: two pulls with the light weight, five with the heavy, three with the light, four with the heavy, three with the light, five with the heavy, four with the light, three with the heavy, five with the light, four with the heavy, and four with the light.

The length of the pull was about 70 cm., the end of which was determined by the weight pulling against a spring fastened to the floor. The fact that the weight moved with only one half the velocity of the arm of the subject, together with the fact that the end of the movement was made by means of a spring, relieved the subject from the blow that otherwise would have been experienced at the end of the movement. Only the first 60 cm. of the movement were recorded, due to the fact that the quick recoil of the spring would have destroyed the last part if the entire movement had been included. The recoil was so rapid that no device we could contrive would stop the vibrations of the fork quickly enough that the return would not blur the true record, and so the smoked surface was removed before the return of the fork. The fact that the movement was 70 cm. in length and the record only 60 cm. made this possible.

In computing the results the graphic records were divided into twelve five-centimeter divisions and the time for each section recorded as accurately as could be done from the markings of a 100 vs. fork. This gave a measure of the speed of each pull in twelve sections. These twelve scores are given in Table I for each subject and the averages for all the subjects

Table I

Time Scores

Subjects		Sections 1	2	3	4	5	6	7	8	9	10	11	12
A..........	1st L....	3.55	1.93	1.57	1.38	1.26	1.19	1.13	1.07	1.05	1.03	1.02	1.00
	Rem. L..	4.07	2.17	1.75	1.53	1.38	1.27	1.19	1.14	1.10	1.07	1.04	1.02
	1st H....	5.30	2.93	2.40	2.15	1.97	1.84	1.72	1.65	1.59	1.54	1.53	1.52
	Rem. H..	4.83	2.74	2.26	1.98	1.82	1.72	1.63	1.56	1.52	1.48	1.46	1.45
B..........	1st L....	7.10	3.47	2.62	2.13	1.85	1.59	1.36	1.19	1.07	0.98	0.93	0.89
	Rem. L..	6.24	2.72	2.08	1.72	1.48	1.31	1.18	1.07	0.99	0.93	0.89	0.87
	1st H....	5.20	2.57	2.05	1.77	1.62	1.49	1.39	1.30	1.24	1.19	1.14	1.11
	Rem. H..	5.75	3.08	2.38	2.03	1.79	1.60	1.46	1.34	1.25	1.18	1.12	1.07
C..........	1st L....	2.89	1.70	1.44	1.30	1.21	1.13	1.07	1.02	0.97	0.93	0.90	0.88
	Rem. L..	3.24	1.86	1.61	1.43	1.30	1.19	1.10	1.04	0.99	0.94	0.90	0.87
	1st H....	3.64	1.88	1.64	1.46	1.34	1.25	1.18	1.12	1.08	1.05	1.02	1.01
	Rem. H..	3.60	2.06	1.65	1.46	1.34	1.25	1.18	1.12	1.07	1.04	1.01	1.00
D..........	1st L....	7.35	3.16	2.27	1.79	1.50	1.31	1.19	1.11	1.05	1.01	0.97	0.95
	Rem. L..	6.67	2.87	2.07	1.68	1.44	1.28	1.17	1.09	1.03	1.00	0.97	0.95
	1st H....	5.32	2.83	2.19	1.90	1.70	1.58	1.48	1.40	1.34	1.31	1.27	1.25
	Rem. H..	5.85	3.09	2.39	2.03	1.78	1.61	1.48	1.38	1.31	1.25	1.20	1.17
E..........	1st L....	1.85	1.41	1.14	1.00	0.92	0.87	0.82	0.76	0.71	0.68	0.67	0.64
	Rem. L..	1.95	1.41	1.13	1.00	0.90	0.86	0.82	0.79	0.77	0.74	0.72	0.71
	1st H....	2.95	1.78	1.33	1.26	1.21	1.14	1.09	1.06	1.04	1.03	1.03	1.04
	Rem. H..	2.94	1.65	1.31	1.18	1.10	1.05	1.00	0.96	0.93	0.92	0.91	0.90
F..........	1st L....	4.28	2.37	1.91	1.66	1.49	1.35	1.22	1.13	1.05	1.00	0.97	0.95
	Rem. L..	4.22	2.19	1.80	1.55	1.39	1.28	1.18	1.10	1.04	0.99	0.95	0.92
	1st H....	5.40	2.59	2.08	1.80	1.60	1.47	1.36	1.28	1.23	1.15	1.10	1.08
	Rem. H..	4.75	2.49	1.98	1.77	1.64	1.53	1.43	1.35	1.27	1.21	1.16	1.12
G..........	1st L....	2.05	1.42	1.16	1.04	0.96	0.93	0.90	0.88	0.87	0.86	0.84	0.82
	Rem. L..	1.98	1.51	1.22	1.10	1.01	0.97	0.94	0.92	0.90	0.89	0.88	0.87
	1st H....	2.95	1.89	1.59	1.47	1.39	1.34	1.33	1.33	1.34	1.35	1.37	1.40
	Rem. H..	2.92	1.87	1.56	1.47	1.40	1.35	1.32	1.31	1.32	1.33	1.34	1.36
H..........	1st L....	6.77	2.76	2.06	1.72	1.53	1.39	1.29	1.23	1.18	1.15	1.13	1.12
	Rem. L..	6.37	2.78	2.16	1.82	1.65	1.51	1.41	1.34	1.28	1.24	1.20	1.18
	1st H....	7.70	3.66	2.86	2.46	2.22	2.07	1.93	1.88	1.84	1.78	1.76	1.75
	Rem. H..	7.41	3.42	2.73	2.32	2.08	1.95	1.85	1.76	1.69	1.65	1.64	1.63
Av..........	1st L....	4.48	2.28	1.77	1.50	1.34	1.22	1.12	1.05	0.99	0.95	0.93	0.91
$P.E._m$......	....	*.58*	*.13*	*.13*	*.10*	*.08*	*.06*	*.05*	*.04*	*.03*	*.03*	*.02*	*.03*
Av..........	Rem. L..	4.34	2.18	1.73	1.48	1.31	1.21	1.12	1.06	1.01	0.98	0.95	0.93
$P.E._m$......	....	*.47*	*.13*	*.09*	*.07*	*.06*	*.05*	*.04*	*.03*	*.03*	*.03*	*.03*	*.03*
Av..........	1st H....	4.82	2.52	2.02	1.79	1.63	1.53	1.44	1.38	1.34	1.30	1.28	1.27
$P.E._m$......	....	*.36*	*.14*	*.10*	*.09*	*.08*	*.07*	*.06*	*.06*	*.06*	*.06*	*.06*	*.06*
Av..........	Rem. H..	4.76	2.56	2.04	1.78	1.62	1.51	1.42	1.35	1.30	1.26	1.23	1.21
$P.E._m$......	....	*.36*	*.16*	*.12*	*.09*	*.07*	*.06*	*.06*	*.05*	*.05*	*.05*	*.06*	*.06*

given at the bottom of the table. Each subject is given four scores for each of the twelve sections of the pulls. The first score is the average time in hundredths of a second of the first pulls with the light weight immediately after a change of weight had been made. The second score is the average of the remaining pulls with the light weight. The third score is the average in units of hundredths of a second of the heavy weight immediately after a change from the light one. The fourth score is for the remaining heavy pulls.

In Table II are given the average scores of the last pull

Table II

Showing the Effect of the First Change in Weights

Times for Last Light Pull before Change

	Sections											
	1	2	3	4	5	6	7	8	9	10	11	12
Av.......	4.50	2.34	1.87	1.60	1.42	1.27	1.17	1.10	1.05	1.01	0.98	0.96
P.E.$_m$....	.37	.17	.12	.10	.08	.06	.05	.05	.04	.04	.04	.04

Times for First Heavy Pull after Change

	Sections											
	1	2	3	4	5	6	7	8	9	10	11	12
Av.......	5.50	2.68	2.17	1.93	1.77	1.68	1.61	1.56	1.51	1.48	1.46	1.45
P.E.$_m$....	.47	.19	.14	.15	.10	.09	.08	.08	.08	.08	.08	.08

Force for First Heavy Pull after Change

	Sections										
	1–2	2–3	3–4	4–5	5–6	6–7	7–8	8–9	9–10	10–11	11–12
Av.......	25.8	21.3	18.5	17.2	14.4	13.7	13.0	12.6	11.3	9.7	8.7
P.E.$_m$....	2.2	1.8	1.0	1.2	.8	.6	1.4	1.1	.9	.8	.9

with the light weight before any change of weights had taken place, and the first score after the change to the heavy weight was made for the first time. Although the time for the last pull with the light weight before any change in weights occurred is slower than the other average time scores with the light weight it is not significantly so. The first pull with the heavy weight shows a slower time than the other scores with the heavy weight, which is significant throughout the full

extent of the pull (see Fig. 1). It is evident that some form of adjustment took place which did not develop as quickly as the total time of the pull, that is, in not less than 250 sigma. It is possible that this adjustment was one of position. The first heavy weight was a total surprise and as a rule the subject was not sufficiently braced to meet the extra load, consequently he could not bring the extra force to play as readily as in the later pulls when he had a greater or less anticipation of the repetition of such a shock. Several of the subjects stated that after the first pull on the heavy weight they braced themselves more strongly; that is, placed one foot further to the front

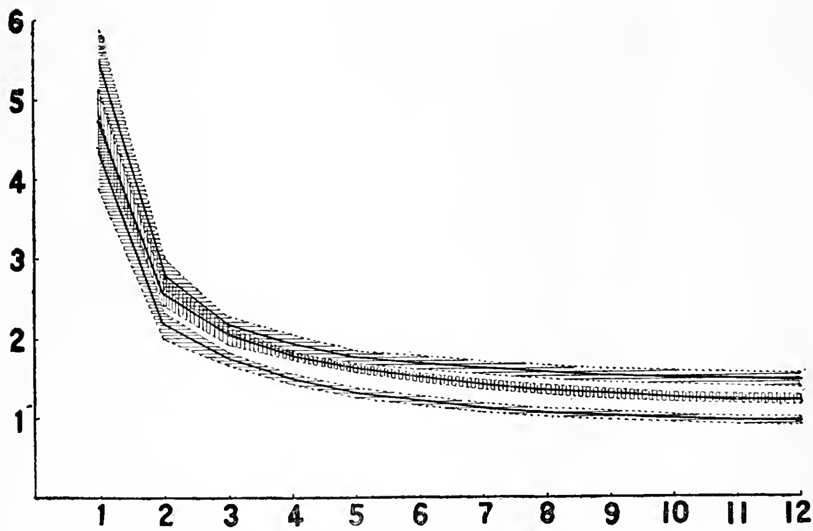


Fig. 1. Graph of the average time scores when the subjects were instructed to use maximum force, with the probable errors at each point represented by the area enclosed within the dotted lines and distinguished by the cross bars. The upper curve represents the first pull with the heavy weight in each experiment, the middle curve represents the pulls with the heavy weight and the bottom curve the pulls with the light weight, the first pull after each change of weights being eliminated.

and took an attitude of greater tension. The average of the first pulls are not reliably different from the average of the remaining pulls, for either the light or heavy weights; showing that after the first shock the changes were met by adequate preparation on the part of the subject.

In Fig. 1 are presented graphically the average time scores, each average being enclosed in a space representing the probable error.[1] The uppermost curve represents the average time taken to pull the heavy weight the first time it was given to each subject. The middle curve is the curve for the heavy weights eliminating the first pull after each change. The lowest is for the light weight with the first pulls after each change eliminated.

In the first part of the pulls the difference in time between the light and heavy weights is not as great as the probable error of the difference. In the fourth section the difference is 3.25 times as great as the probable error of the difference and the reliability increases to the end of the pull when the difference is 4.17 times the probable error of the difference. This shows that in the first part of the pulls the heavy and light weights were lifted with nearly the same speed but that as the pull progressed the heavy weight was pulled more slowly than the light one.

Still further light is thrown on the nature of the adjustments which occurred upon the change of weights by the force scores; but before discussing their significance it may be well to explain how they were derived. For each of the twelve sections of the pull the average velocity was found by dividing the extent by the time; that is, each score in Table I was divided into 5, the extent of each section. The acceleration was then found by subtracting the average velocity of one five-centimeter space from the average velocity of the succeeding five-centimeter space and the remainder divided by the average time for the two spaces. Having found the acceleration the force was next calculated, it being the product of the acceleration plus the acceleration of gravity (980 cm. per sec.) and the mass. In brief the formulas are:

$$v = \frac{s}{t}, \qquad a = \frac{V_2 - V_1}{t}, \qquad F = M(a + 980),$$

in which s = extent, t = time, a = acceleration, V_1 = initial velocity, V_2 = final velocity, F = force in dynes, and M = mass.

[1] The writer is indebted for this method of presentation of the probable error to a suggestion of Professor Cattell.

Table III

Force Scores

Subjects	Sections 1–2	2–3	3–4	4–5	5–6	6–7	7–8	8–9	9–10	10–11	11–12
A, 1st L.	12.90	10.60	9.65	8.66	7.36	6.79	5.71	4.69	4.50	4.30	3.84
Rem. L.	10.75	9.37	8.35	8.42	8.07	7.55	6.15	5.91	5.54	5.39	4.52
1st H.	22.18	18.38	15.82	15.70	15.37	15.47	13.60	12.89	12.58	9.13	8.62
Rem. H.	23.72	19.73	18.62	16.38	15.07	15.03	13.44	12.15	11.75	11.31	8.16
B, 1st L.	5.81	6.07	6.79	6.92	8.35	11.50	12.15	13.15	12.35	9.56	8.84
Rem. L.	8.05	7.99	9.20	9.26	10.09	10.80	11.48	11.60	10.77	8.84	6.00
1st H.	28.40	24.39	22.07	19.97	20.58	21.09	21.44	19.25	19.11	18.96	15.93
Rem. H.	21.00	20.98	20.66	22.30	21.44	22.83	23.70	24.39	22.95	22.45	22.49
C, 1st L.	15.25	10.80	8.98	8.22	8.24	8.15	8.01	8.03	8.04	7.48	5.68
Rem. L.	20.20	8.01	8.50	8.81	9.41	9.65	8.54	8.64	8.78	8.76	7.65
1st H.	43.80	24.80	26.15	24.20	24.39	22.99	23.15	19.65	17.82	17.35	11.45
Rem. H.	36.20	32.70	27.11	24.20	24.39	22.99	23.15	21.79	18.66	18.22	11.47
D, 1st L.	6.57	8.86	9.46	10.25	10.90	10.00	8.72	8.50	6.88	6.82	4.92
Rem. L.	7.46	8.90	9.95	10.05	10.10	9.95	9.07	8.37	6.00	5.86	5.69
1st H.	23.42	23.07	21.29	21.44	22.10	18.26	18.42	16.12	14.64	13.02	11.92
Rem. H.	20.82	20.94	20.33	21.09	21.91	20.66	21.21	19.15	18.54	18.39	14.17
E, 1st L.	15.10	18.25	16.50	13.55	12.20	12.21	16.90	17.65	13.60	11.42	10.21
Rem. L.	16.75	19.10	15.65	16.55	10.15	9.96	9.65	9.63	8.25	8.08	6.49
1st H.	43.80	55.00	20.20	18.93	23.50	22.95	16.28	14.26	8.69	7.61	7.56
Rem. H.	52.70	48.50	34.50	28.00	24.95	24.31	23.46	22.42	13.51	12.70	11.88
F, 1st L.	9.21	8.32	7.73	7.35	8.74	9.99	9.44	9.79	10.00	6.35	5.18
Rem. L.	10.75	8.52	8.80	8.38	8.42	8.74	9.02	8.54	8.16	7.70	7.07
1st H.	27.23	22.91	23.23	23.15	21.80	22.45	21.13	18.14	25.95	20.74	14.00
Rem. H.	28.20	25.02	20.47	18.10	17.40	18.15	19.89	20.66	20.74	18.77	17.82
G, 1st L.	17.60	17.30	13.45	12.15	6.77	6.66	6.23	4.89	4.36	6.40	6.50
Rem. L.	13.55	16.30	11.43	11.88	7.32	6.73	5.54	5.10	4.57	4.32	4.06
1st H.	37.75	29.90	19.81	19.57	14.44	9.36	7.61	6.44	5.32	4.71	2.64
Rem. H.	38.40	31.60	17.86	16.81	14.95	12.23	9.38	5.84	5.82	5.80	4.73
G, 1st L.	7.95	8.69	8.32	7.80	8.08	7.30	6.25	5.83	4.69	3.89	3.25
Rem. L.	8.38	7.23	7.68	6.46	6.56	6.57	5.58	5.56	5.10	4.98	3.81
1st H.	17.23	16.89	15.51	14.91	14.13	13.82	10.47	10.12	11.03	10.14	8.50
Rem. H.	18.38	15.89	16.42	15.66	13.78	13.74	13.20	12.56	11.35	8.55	8.56
Av. 1st L.	11.29	11.11	10.11	9.36	8.83	9.07	9.18	9.07	8.05	7.03	6.05
$P.E._m$	*1.17*	*.99*	*.73*	*.51*	*.41*	*.55*	*.82*	*.93*	*.88*	*.55*	*.55*
Av. Rem. L.	11.99	10.68	9.95	9.98	8.76	8.74	8.13	7.92	7.15	6.74	5.66
$P.E._m$	*1.09*	*1.05*	*.61*	*.64*	*.35*	*.40*	*.53*	*.61*	*.55*	*.48*	*.36*
Av. 1st H.	30.48	26.92	20.51	19.73	19.54	18.30	16.51	14.61	14.39	12.73	10.08
$P.E._m$	*2.54*	*2.32*	*.80*	*.74*	*1.10*	*1.22*	*1.36*	*1.10*	*1.50*	*1.44*	*1.97*
Av. Rem. H.	29.93	26.92	22.00	20.32	19.24	18.74	18.43	17.37	15.41	14.52	12.41
$P.E._m$	*2.81*	*2.40*	*1.32*	*1.07*	*1.18*	*1.19*	*1.44*	*1.61*	*1.44*	*1.48*	*1.29*

In Table III the force records are arranged to correspond to the time scores in Table I. The first column gives the force scores (in units of 1,000,000 dynes) computed from the times of the first and second sections of the pulls; that is, from columns one and two in Table I. The second column in Table III represents the force scores computed from columns two and three of Table I., etc.

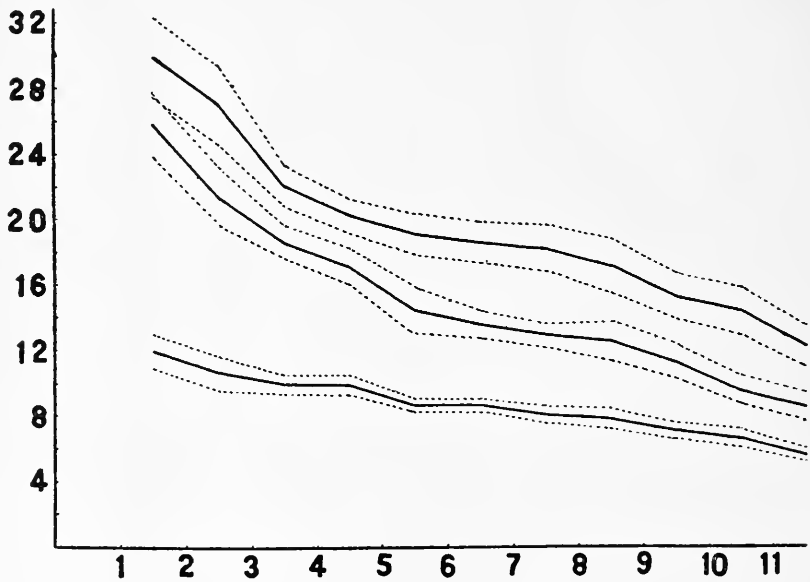


Fig. 2. Graph of the average force scores when the subjects were instructed to use maximum force, with the probable errors at each point represented by the area enclosed within the dotted lines. The middle curve represents the first pull with the heavy weight in each experiment, the upper curve represents the pulls with the heavy weight and the bottom curve the pulls with the light weight, the first pull after each change of weights being eliminated.

The force records (Table III and Fig. 2) show clearly the significance of the fact pointed out in the time scores, namely that the first part of the pulls with either light or heavy weight tend to have the same time values. This of course means that a much greater amount of force had to be used in pulling the heavy weight than was used in pulling the light one. Toward the end of the pulls the difference in force is less marked although it is a highly reliable difference, being five times the probable error of the difference.

The striking fact here is that this large difference in the force used in pulling two greatly differing weights is present at the very beginning of the first pull after the change. That is, after one has been pulling a weight of 2,440 grams with what he supposes to be the maximum force that he is able to exert, when unexpectedly a weight of 7,770 grams is substituted for the lighter one, his force at the very beginning of the pull is on the average 2.5 times as great as the supposedly maximum force previously used.

The time included in the calculation of the force scores in the first column embraces the time it took the subject to pull the weight the first 7.5 cm., or the first time score and half the second. This time ranges from 0.025 sec. to 0.091 sec. with an average of 0.054 sec., much shorter in every case than the simple reaction time, which under the most favorable circumstances can scarcely be reduced to 0.100 sec. (7). In much less time than it takes one to make a simple reaction an adjustment in force can be made when an unexpectedly heavy or light weight is raised. It is evident that this adjustment is either of a reflex nature, or it is something even more elementary.

In every case the greatest amount of force is exerted at the beginning of the pull, the difference between the beginning and end being greater with the heavy weight. Probably this can be explained by an analysis of the manner of pulling in the light of the data already presented. Each pull was started from a position which allowed the rope to sag, the subject setting himself for his maximum pull, which doubtless was to pull as quickly as possible; hence the speed at the initial part of the pull was about the same in all cases. After the first impulse the light weight continued at a rapid rate with little continued force while the heavy one would require more force to keep it going at the pace given by the first impulse. With jerks at the beginning which would start both light and heavy weights at the same rate the light weight would end the movement at a greater rate of speed than the heavy one.

While we have shown that one's execution of what he regards as his maximum force is determined by the amount

of resistance opposed to the movement involved and while the quickness with which the force adjustment is made indicates that it is of a reflex or some simpler nature, it is not so certain that under more favorable circumstances one could not make a proper time adjustment to a change in weight and keep the resultant physical force the same. In order to see with what degree of accuracy such an adjustment could be made, another experiment was performed in which the subjects were told just what would take place.

The directions given were as follows: "The object of this experiment is to test your ability to use the same force in pulling different weights. There will be four different weights and you will be given the next to the heaviest first. You will always be notified when a weight is changed and given the privilege of feeling how heavy it is before making the pulls. Three pulls will be made between changes and you will be asked to judge in which of the three you think you used the nearest to the same amount of force as with the preceding weight. Make a pull each time you are given the buzzer signal to do so, and start each pull from a position which will allow the rope to sag. *Remember* it is the force you use that you are to keep the same regardless of the time it takes you to make the pull."

Four weights were used instead of only two as in the first experiment of the paper. These weights were 15,540; 12,180; 9,300; and 4,880 grams; which, due to the fact already stated that they moved only one half the distance of the recording carriage, were treated in the tabulations as weights of 7,770; 6,090; 4,650; and 2,440 grams. Each weight was given at three different times during the experimental period and at each presentation was given three pulls. After the first three pulls the subject was asked which one of the three he judged nearest to what he wished to keep as a standard. Thereafter after each three pulls he was asked to judge which of the three came nearest to the standard he had so chosen. No tabulation of the accuracy of these judgments is presented since the poorest record was as often chosen as the best, and all the subjects said that they were merely guessing and had not the least idea how nearly they were coming to the same force.

If we let I. represent the 2,440 gram weight, II. the 4,650, III. the 6,090, and IV. the 7,770, the order in which they were given was III., I., II., IV., I., II., III., IV., III., IV., I., and II. In the calculations only the last eight presentations were used. The first four were regarded as practice pulls, for although the subjects were told that next to the heaviest was to be given first they could not adjust themselves to this information. Two of the subjects when they came to the heaviest weight for the first time said that if they used the same force that they had been using they could not pull it at all. They were then told to pull IV. to suit themselves and to try to use this new standard throughout the remainder of the experiment.

It may seem from this that the subjects had some idea of what it meant to use the same force. When asked what they meant by changing the amount of force in order to pull the heaviest weight, they stated that they had to change their posture, that they had been controlling their former pulls by setting their feet a certain distance apart and pulling so that with the same set of their musculature they were neither drawn forward nor moved backwards when they pulled. When the heavy one came they felt that this posture was not sufficiently stable and experienced a pull forward, hence the necessity for readjustment of position. We may state here that the weights were of such a range that it was physiologically possible to adjust the time of pulling so that each could be pulled with the same force.

Eight subjects were used, two of whom had served in the first experiment. In this experiment two of the subjects were women, *J* and *L*. On reading the directions most of them seemed bewildered and their remarks and attitude throughout the experiment showed that the task of keeping force the same with different weights was one to which they were unaccustomed. Most of them reassured themselves before they began by asking if the directions did not mean that they must pull the heavy weights slower than the light ones. The fact that they required such information is an indication that they had a very crude idea of force as such. They simply

Table IV

Time Scores

Subjects	Wt.	Sections 1	2	3	4	5	6	7	8	9	10	11	12
I.	I.	8.2	3.4	2.7	2.3	2.0	1.9	1.9	1.9	1.9	2.0	2.1	2.2
	II.	7.6	3.9	2.9	2.6	2.3	2.1	2.2	2.2	2.2	2.4	2.7	2.8
	III.	10.8	5.3	3.9	3.4	3.2	2.9	3.0	3.0	3.0	3.2	3.6	4.2
	IV.	13.8	8.6	7.4	6.8	6.4	6.1	5.9	5.6	5.4	5.7	6.0	7.1
B.	I.	8.7	4.1	3.2	2.7	2.4	2.2	2.1	2.0	1.9	1.9	1.87	1.83
	II.	8.8	4.7	3.5	3.0	2.7	2.5	2.5	2.4	2.4	2.4	2.4	2.6
	III.	9.5	5.1	3.8	3.3	3.1	2.8	2.9	2.9	2.9	3.0	3.1	3.4
	IV.	10.3	5.4	4.2	3.6	3.4	3.1	3.2	3.2	3.2	3.4	3.5	3.9
J.	I.	15.3	7.3	5.8	4.8	4.4	4.0	4.0	4.0	4.3	5.1	6.8	9.8
	II.	30.3	12.3	9.6	8.4	8.0	7.8	7.7	8.1	8.7	9.8	10.9	12.1
	III.	27.5	14.6	12.6	13.0	12.1	12.4	11.8	10.7	11.6	12.0	13.6	16.8
	IV.	31.0	17.0	13.3	13.9	14.2	14.3	15.3	13.9	14.3	15.8	14.3	17.0
G.	I.	3.37	2.21	1.87	1.75	1.62	1.58	1.58	1.54	1.54	1.58	1.62	1.67
	II.	3.75	2.37	2.08	1.87	1.83	1.75	1.75	1.79	1.87	1.96	2.04	2.17
	III.	4.75	2.75	2.46	2.21	2.08	2.04	2.00	2.00	2.00	2.10	2.20	2.40
	IV.	4.62	2.96	2.62	2.37	2.25	2.21	2.17	2.25	2.37	2.58	3.00	3.00
K.	I.	6.5	3.08	2.46	2.12	1.92	1.79	1.58	1.5	1.49	1.42	1.42	1.46
	II.	10.3	5.6	4.4	3.9	3.8	3.8	3.7	3.7	3.8	3.8	3.8	4.0
	III.	10.3	6.5	5.6	5.0	5.0	5.1	5.4	5.5	5.3	5.0	5.1	5.4
	IV.	13.5	9.3	8.6	7.7	7.8	8.3	12.0	8.5	7.8	8.0	8.8	12.3
L.	I.	4.2	2.6	2.0	1.8	1.6	1.5	1.4	1.3	1.27	1.24	1.22	1.2
	II.	6.3	3.6	2.9	2.5	2.2	2.0	1.9	1.8	1.7	1.7	1.75	1.79
	III.	6.5	3.8	3.2	2.7	2.4	2.3	2.25	2.2	2.1	2.1	2.12	2.17
	IV.	8.0	4.7	3.8	3.5	3.1	3.0	2.8	2.8	2.8	2.8	3.0	3.2
M.	I.	3.8	2.3	2.0	1.8	1.6	1.5	1.5	1.5	1.5	1.5	1.54	1.58
	II.	4.5	2.8	2.3	2.0	1.8	1.7	1.7	1.6	1.6	1.6	1.65	1.7
	III.	5.6	3.4	2.7	2.4	2.2	2.1	2.0	2.0	2.0	2.1	2.2	2.2
	IV.	6.0	3.6	3.0	2.5	2.3	2.2	2.1	2.05	2.0	2.04	2.09	2.15
N.	I.	14.2	16.0	16.3	15.2	14.4	14.1	11.6	11.4	10.9	9.5	8.5	8.2
	II.	26.3	15.4	14.3	14.3	13.1	13.0	11.9	12.0	13.2	13.0	11.4	9.4
	III.	31.7	18.1	15.2	15.7	16.5	14.8	14.7	13.5	14.1	14.0	13.8	13.5
	IV.	25.7	16.1	13.1	14.0	14.3	13.7	12.2	12.6	13.2	11.5	9.9	10.6
Av.	I.	8.0	5.13	4.53	4.06	3.74	3.70	3.21	3.14	3.10	3.03	3.13	3.50
$P.E._m$ . . .		*1.1*	*.9*	*.9*	*.9*	*.8*	*.8*	*.8*	*.8*	*.7*	*.6*	*.8*	*.9*
Av.	II.	12.2	6.27	5.2	4.82	4.47	4.32	4.17	4.20	4.42	4.56	4.57	4.57
$P.E._m$ . . .		*2.4*	*1.0*	*1.0*	*1.0*	*.9*	*1.0*	*.8*	*.9*	*1.0*	*1.0*	*1.0*	*1.0*
Av.	III.	13.3	7.42	6.17	5.95	5.7	5.55	5.55	5.23	5.37	5.42	5.71	6.26
$P.E._m$ . . .		*2.4*	*1.1*	*1.2*	*1.1*	*1.3*	*1.2*	*1.2*	*1.1*	*1.1*	*1.1*	*1.1*	*1.1*
Av.	IV.	14.1	8.45	7.00	6.80	6.7	6.61	6.95	6.37	6.38	6.47	6.32	7.42
$P.E._m$ . . .		*2.1*	*1.3*	*1.5*	*1.1*	*1.2*	*1.3*	*1.4*	*1.2*	*1.2*	*1.2*	*1.2*	*1.3*

tried from what they knew of the laws of dynamics to compensate for a heavy weight by a slower speed.

In this case the scores were calculated as before, that is the total extent of the record, 60 cm., was divided into twelve sections and the time taken for each section. In Table IV four lines of scores are given for each subject; I. the times in hundredths of a second for each section with the 2,440 weight, II. for 4,650, III. for 6,090, and IV. for 7,770. From the time scores the force scores in Table V were derived by means of the formulas given above. The force unit is again 1,000,000 dynes. The averages of these two tables are graphed in Figs. 3 and 4, with the areas representing the range of their probable errors embraced by dotted lines.

The first thing evident is that the time scores do not coincide as nearly as they did when the instructions were to use maximum force throughout. This is no doubt the result of the deliberate attempt to change the time to compensate for changes in the weight. Taken by themselves the time scores might seem to indicate the ability to make such an adjustment, but a glance at the force records will show that the adjustment is at best very poor. The difference in the first time scores between the light and heavy weights is 2.5 times the probable error of the difference, the difference in the last time scores 1.6 times the probable error of the difference. The difference between the lightest and heaviest is for the first force score 6 times the probable error of the difference, for the last force score 6.05 times the probable error of the difference. The relative reliability of the time and force records can be seen clearly by the relative overlapping of the probable error areas in Figs. 3 and 4.

In this experiment the pulls were of a different nature from those when the subject tried to use maximum force; they were less violent and the force was sustained more throughout the extent of the movement. Even here however more force was used at the beginning than at the end of the pull. The time curves show a rapid increase in rate at the beginning and then a fairly even rate through the remainder of the pull with a slight indication of a diminution of speed at the end.

Table V

Force Scores

Subjects	Wts.	Sections 1–2	2–3	3–4	4–5	5–6	6–7	7–8	8–9	9–10	10–11	11–12
I........	I.	6.00	4.99	5.55	6.07	4.01	2.39	2.39	2.39	0.76	0.84	1.25
	II.	9.23	10.62	7.90	9.08	8.82	2.18	4.56	4.56	0.81	0.29	2.33
	III.	9.00	10.45	9.11	7.64	9.17	4.79	5.97	5.97	3.87	2.90	2.86
	IV.	9.13	8.52	8.28	8.14	8.12	8.96	8.24	8.08	6.93	7.04	5.94
B........	I.	4.84	4.65	5.76	4.62	4.35	3.63	3.82	4.01	2.39	2.90	3.17
	II.	7.83	8.70	7.97	7.55	7.22	4.56	5.68	4.56	4.56	4.56	1.55
	III.	9.62	10.53	9.44	7.83	9.53	4.70	5.97	5.97	4.79	4.90	3.29
	IV.	11.95	11.88	11.54	9.41	11.04	6.31	7.61	7.61	5.49	6.48	4.52
J........	I.	3.16	3.15	3.22	2.94	3.05	2.39	2.39	1.89	1.43	1.39	1.73
	II.	5.08	5.04	4.93	4.71	4.65	4.60	4.34	4.32	4.21	4.32	4.39
	III.	6.43	6.22	5.92	6.11	5.92	6.07	6.22	5.77	5.90	5.73	5.70
	IV.	8.04	8.04	7.51	7.57	7.58	7.49	7.78	7.55	7.43	7.78	7.34
G........	I.	8.15	7.28	4.82	4.87	3.54	2.39	3.72	2.39	1.07	1.24	1.06
	II.	16.41	10.62	10.90	6.07	7.92	4.56	2.72	1.51	1.52	2.35	1.13
	III.	18.38	10.98	11.96	9.93	7.45	7.48	5.97	5.97	2.41	2.86	0.94
	IV.	19.65	14.74	12.66	10.64	10.92	7.61	4.34	4.63	2.56	0.58	0.19
K.......	I.	6.72	6.00	5.86	5.34	4.87	5.01	5.14	2.73	5.25	2.39	0.69
	II.	6.94	6.81	6.20	4.94	4.56	5.02	4.56	4.09	4.56	4.56	3.81
	III.	8.03	7.21	7.20	5.97	5.73	5.34	5.77	6.37	6.64	5.73	5.34
	IV.	8.00	8.00	8.26	7.54	7.23	6.20	8.90	8.12	7.44	7.09	6.41
L.......	I.	7.65	8.54	5.93	7.42	5.66	6.29	7.30	3.85	4.32	3.79	3.59
	II.	10.13	9.35	9.29	9.87	9.63	7.66	8.21	8.95	4.56	2.40	2.67
	III.	12.39	10.25	11.96	11.56	8.31	7.17	7.35	9.08	5.97	5.39	3.42
	IV.	12.97	12.23	10.06	11.95	8.98	10.78	7.61	7.61	7.61	4.46	4.92
M.......	I.	5.83	6.07	5.92	7.42	5.66	2.39	2.39	2.39	2.39	0.93	1.29
	II.	12.14	11.67	11.57	11.30	8.93	4.56	9.76	4.56	4.56	0.97	1.09
	III.	13.76	13.58	11.53	10.86	9.09	9.53	5.97	5.97	2.40	2.86	5.97
	IV.	16.58	13.97	17.00	13.28	10.90	11.56	9.48	10.29	5.78	5.36	4.81
N.......	I.	2.33	2.38	2.42	2.41	2.40	2.54	2.40	2.43	2.55	2.53	2.45
	II.	4.83	4.63	4.56	4.67	4.57	4.68	4.55	4.41	4.58	4.74	4.97
	III.	6.25	6.17	5.94	5.91	6.11	5.97	6.09	6.03	5.97	5.99	6.00
	IV.	8.04	7.98	7.45	7.57	7.68	7.87	7.53	7.72	7.96	7.96	7.86
Av.......	I.	5.58	5.38	4.81	5.14	4.19	3.38	3.69	2.76	2.52	2.00	1.90
P.E.$_m$....		*.47*	*.47*	*.46*	*.43*	*.28*	*.36*	*.39*	*.18*	*.34*	*.27*	*.26*
Av.......	II.	9.07	8.43	7.91	7.27	7.04	4.73	5.55	4.62	3.67	3.02	2.74
P.E.$_m$....		*.79*	*.66*	*.61*	*.65*	*.55*	*.20*	*.52*	*.32*	*.30*	*.46*	*.37*
Av.......	III.	10.48	9.42	8.12	8.23	7.66	6.38	6.16	6.39	4.74	4.54	4.19
P.E.$_m$....		*.83*	*.65*	*.40*	*.57*	*.41*	*.38*	*.09*	*.20*	*.41*	*.38*	*.46*
Av.......	IV.	11.79	10.67	10.34	9.51	9.05	8.35	7.68	7.70	6.40	5.59	5.25
P.E.$_m$....		*.93*	*.76*	*.75*	*.55*	*.43*	*.40*	*.28*	*.26*	*.40*	*.63*	*.49*

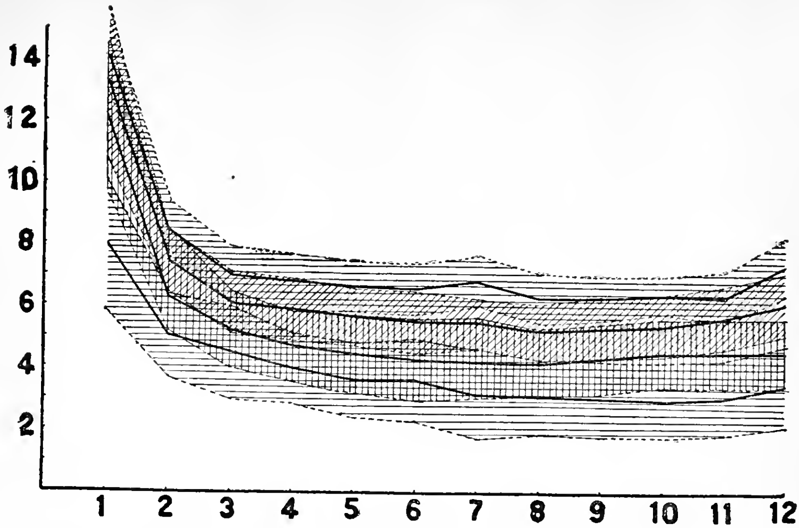


Fig. 3. Graph of the average time scores in the experiment where the subjects tried to work with equal force with different weights. The probable errors at each point are represented by the area enclosed within the dotted lines and distinguished by the cross bars. From top to bottom the curves represent the times with weights of 7,770; 6,090; 4,650, and 2,440 grams.

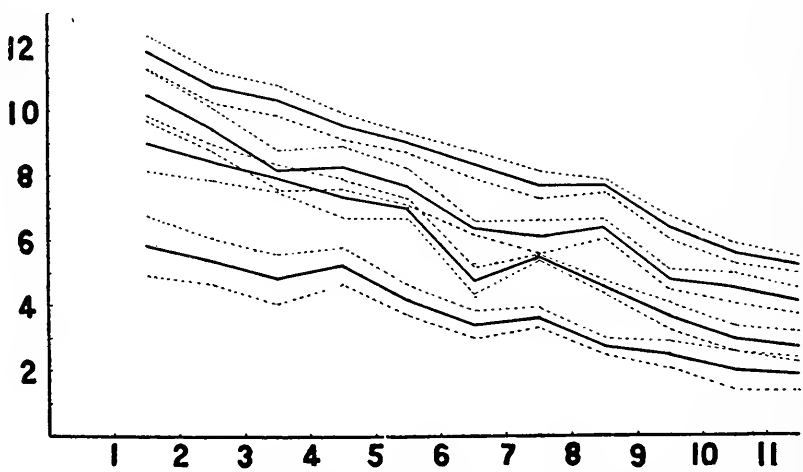


Fig. 4. Graph of the average force scores in the experiments where the subjects tried to work with equal force with different weights. The probable errors at each point are represented by the area enclosed within the dotted lines. From top to bottom the curves represent the force used with weights of 7,770; 6,090; 4,650, and 2,440 grams.

This is evidence that one has no idea how to accelerate a movement so that the force would be constant throughout. Constant force throughout a movement would require a uniform increase in rate for each unit of space. In such a case the time curve would be a straight 45° line in which for each unit of extent there is an equal increase in speed. Instead of that the subjects tended to maintain uniform velocity. In the first part of the movement they quickly caught the velocity they judged proper for that weight and then maintained that velocity at least approximately throughout the pull. They certainly had no idea of the control of force in the sense in which it is regarded in physics. They regarded uniform speed throughout a pull as uniform force.

Conclusions

None of the work done on the judgment or control of force has eliminated both the factors of extent and time or varied the conditions so that the control of the force of a movement as an independent entity could be determined. It is possible that in some of the experiments the controlling factor was time and in others extent, while in some both factors could feature as far as the controls used in the experimental procedure were concerned. We therefore do not regard the data we have obtained as contradictory to that derived from previous experimentation, we simply believe that the work done has not been of the nature to answer the question as to whether one can control force independently of extent of time. Even our experiments do not show conclusively that one cannot control force at all. What it does show is that one has not the ability to adjust the speed of a movement so that with different loads the movement can be made with the same physical force. The *implication* from this would be that one cannot control the amount of force he is using, and that what control he does use is based on some other factor. To test the validity of this implication the writer is working on a different series of experiments which he hopes will show the relative accuracy of the control of force, extent, and time.

The speed with which force is adjusted to the size of the load shows that it is a very elementary process. The question is still left whether it is a local muscular phenomenon or a reflex. If it is a reflex it is a very rapid one, for in some cases the adjustment took place in less than 0.025 sec. It is very likely that it may be a local phenomenon of the muscle itself. Physiologists have found a similar phenomenon in the nerve muscle preparation of the frog. Luciani (6) says: "According to the observations originally made by Fick, and afterwards confirmed by others, when the weight applied to the muscle is not great, and particularly when an elastic resistance is opposed to the muscle, so that its tension increases constantly during contraction, the shortening is greater when the weight and the initial resistance are increased. This paradoxical phenomenon is a specific property of the substance of living muscle, and shows that the sudden pull of the muscle and increase of tension during shortening act as a stimulus on the contractile substance, and increase the effect of the electrical stimulation."

While the quickness of the adjustment in the procedure where maximal force was required may be explained as a local muscular phenomenon it does not show that this is the case where the subjects chose the speed of their pulls and were aware of the size of the weight before they pulled. It is not hard to see, however, how it could be of such a nature even in such a case. If it was controlled by anything above a reflex center the adjustment could certainly be made with greater accuracy than our records show. In short, when the extent is kept constant and the load changed one judges his force by the time of his pull; when he is told that the weights are changed, knowing that his pull should be slower with the heavier, he makes a time correction but it is a very, very crude one.

References

1. d'Allonnes, Révault. Expériences sur l'effort volontaire dans l'évaluation des poids. *L'Année Psychologique*, 1902, 8, 299–325.
2. Claparède, Ed. La vitesse des soulèvements lors des illusions de poids. *L'Année Psychologique*, 1901, 7, 646–8.

3. Fullerton and Cattell. On the Perception of Small Differences, with Special Reference to the Extent, Force, and Time of Movement. Univ. of Penna. Philosophical Series, 1892, No. 2.
4. Hollingworth, H. L. The Inaccuracy of Movement. *Archives of Psychol.*, 1909, No. 13.
5. Jacobi, C. Untersuchungen über den Muskelsinn. *Archiv für exp. Pathol. und Pharmakol.*, 1893, **32**, 49–100. See discussion of his work by Woodworth in *Le Mouvement*, p. 35.
6. Luciani, L. Human Physiology. Trans. by F. A. Welby, 1915, Macmillan, Vol. III, pp. 13, 15, and 46.
7. Ladd and Woodworth. Physiological Psychology, 1911, p. 489.
8. Martin and Müller. Die Unterschiedsempfindlichkeit. Leipzig, 1899.
9. Morgan, J. J. B. The Overcoming of Distraction and Other Resistances. *Archives of Psychol.*, 1916, No. 35. Chap. VII.
10. Müller and Schumann. Ueber d. psychol. Grundlagen d. Vergleichung gehobner Gewichte. *Archiv f. d. ges. Physiol.*, 1889, **45**, 37.
11. Steffens, Laura. Ueber die motorische Einstellung. *Ztsch. f. Psychol. u. Physiol. d. Sinnesorgane*, 1900, **22**, 241.
12. Woodworth, R. S. The Accuracy of Voluntary Movement. *Psychol. Mon.*, 1899, No. 13.
13. Woodworth, R. S. *Le Mouvement*, Paris, 1903.
14. Woodworth, R. S. On the Voluntary Control of the Force of Movement. *Psychol. Rev.*, 1901, **8**, 450–9.

THE JOHNS HOPKINS CHRONOSCOPE

BY KNIGHT DUNLAP

The chronoscope illustrated herewith is the second model, of which nine instruments have been made. The first chronoscope, which has been in operation for over a year, differs from these later ones in the details of the clutch mechanism, but operates in essentially the same way. Twenty-five instruments of the third model, which has some improvements in details of construction, are now being constructed.

The chronoscope is fundamentally a 10-pole synchronous motor (Fig. 1) constructed with extreme care. The shaft bearings are adjustable and replaceable. The motor will run on direct current, interrupted by a tuning fork, from 15 to 100 times per second, or will run without interruption on an alternating current at frequencies of from below 10 to above 90.

The rotation rate of the armature when run on direct current is one tenth the number of interruptions per second, *e. g.*, with a 50 vibration fork, the armature makes five rotations per second. On the alternating current the rotations are one fifth of the frequencies per second, *e. g.*, on the 60-cycle current, the armature makes twelve rotations per second.

The motor operates continuously with very little noise. If it is desirable, the electric fork may be put in another room or may be enclosed in a padded box with a thermostat to keep the temperature uniform. Temperature changes have no effect on the chronoscope mechanism itself: the only part of the system which can be influenced by temperature changes is the tuning fork.

The clutch mechanism, through which the hand is alternately engaged with the motor shaft and braked to a standstill, consists of two simple electromagnets (K' and K''), the core of one (K'') being fastened to the armature shaft and rotating with it. The core of the other (K') may be

rotated by turning the milled ring (F), but is normally at rest. The windings of both magnets are fixed in position so that there is no traveling or moving electric contact. When the current is off both magnets the hand, J, on the dial plate may be set back to zero by turning the milled head H, attached to the large gear wheel G (Fig. 2). When the current is on the forward magnet (K', Fig. 1) the hand may be moved by turning the milled ring, F. When the current is on both

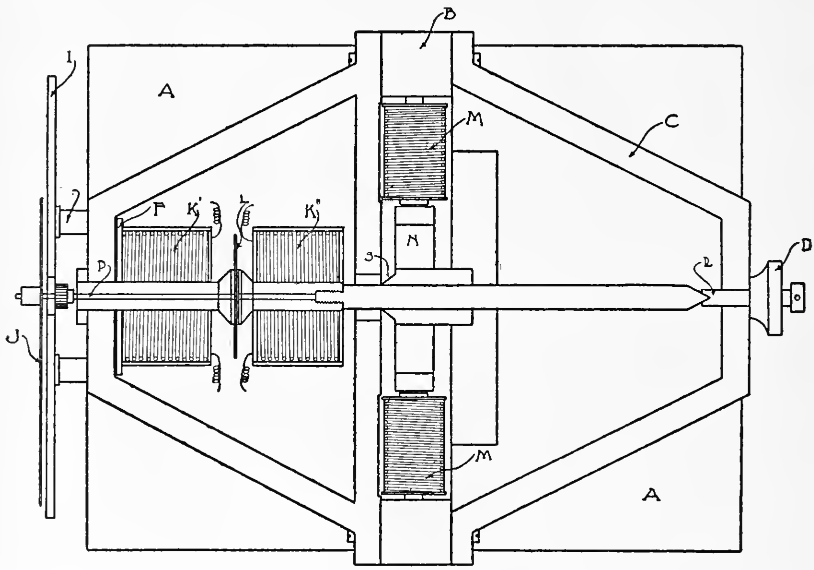


FIG. 1.

magnets, the armature disc, L, carried by the hand shaft, P, will stay against either magnet core. It may be, therefore, moved over from one to the other by breaking first the one circuit and then the other, just as is the case with my method of using the Hipp chronoscope without armature springs. In this way the effect of variations in the electric current actuating the clutch mechanism are entirely avoided. There is a trifle of slip in starting and stopping with this new mechanism, but the error amounts to less than one *sigma* on the average, which is a negligible magnitude for all reaction time work.

The dial, *I* (Fig. 3), is five inches in diameter, and has in the upper portion an aperture, *T*, through which the large gear wheel, *G* (Fig. 2), is visible. This wheel carries numerals, legible through the aperture, which indicate complete rotations of the index hand.

The advantages of the Johns Hopkins Chronoscope are: first, that it may be operated with practically no noise; second, that it requires no winding, since it will run for several hours at a time without further adjustment of the fork contacts; third, that the dial is large and easily read; and fourth, that the hand is readily set back to zero, avoiding the necessity of subtraction.

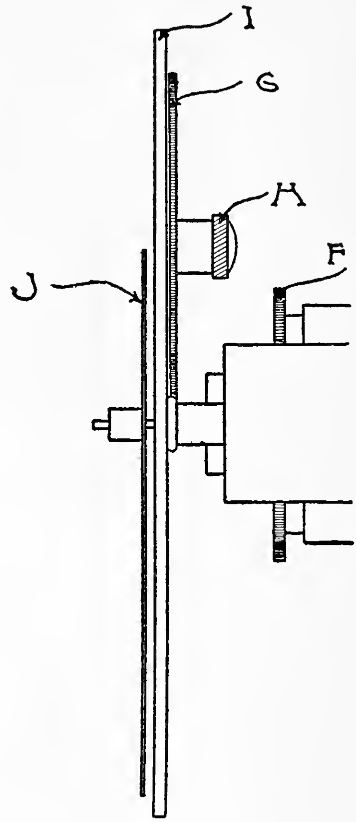


Fig. 2.

The synchronous motor was invented independently by Lord Rayleigh (who published an account of his "phonic wheel" in 1879), and by La Cour (somewhat earlier than Rayleigh). Numerous applications to scientific and commercial purposes have since been made, among them the use in a printing telegraph system, patented by Professor Rowland of the Johns Hopkins University in 1896. My first attempts at construction of a synchronous motor for driving a time machine (in 1908–10) were not completely successful. In 1914 Dr. Lorenz of the Nela Research Laboratory kindly loaned me a large synchronous motor which he had reconstructed from an ordinary A.C. motor, and after using this and discussing the matter with Dr. Lorenz I designed the motor which is the basis of the Johns Hopkins chronoscope. This motor embraces Lorenz's plan of a multipolar field (thus departing

from the Rayleigh, Rowland and Bull models) but in the details is original. My indebtedness to Dr. Lorenz (who also

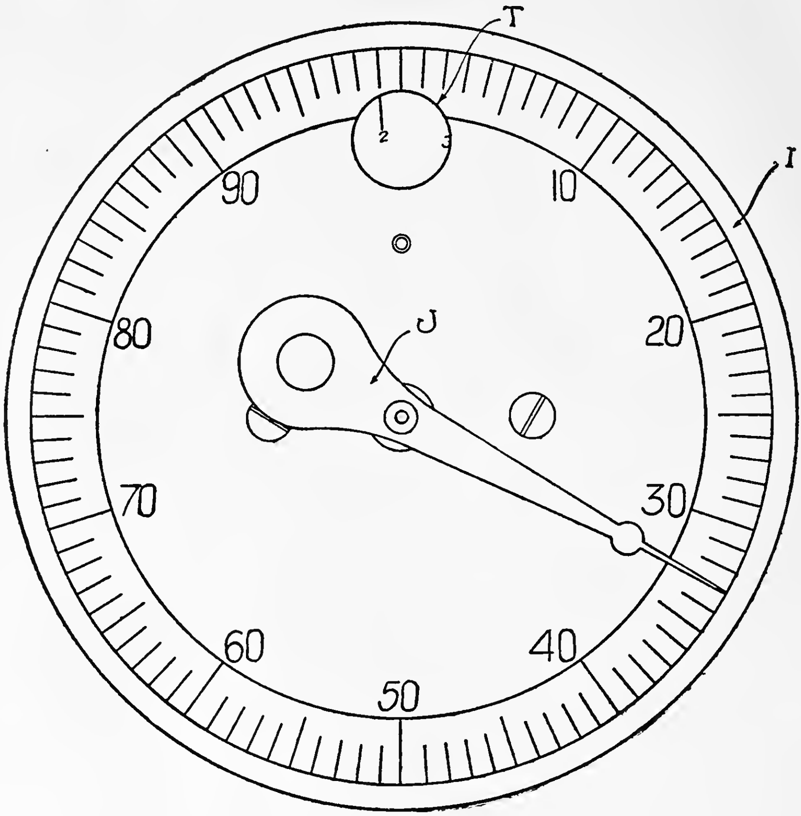


Fig. 3.

applied the synchronous motor to the Seashore tonoscope) is gratefully acknowledged.

Journal of

Experimental Psychology

VOL. II, No. 3. AUGUST, 1917

A NEW CHRONOSCOPE AND FALL APPARATUS

BY PAUL E. KLOPSTEG

University of Minnesota[1]

Certain well-defined principles in physics which are, perhaps, not very familiar to the worker whose interest lies in other fields of science, lend themselves admirably to the measurement of time intervals of the order of magnitude in which the psychologist is especially interested. In this paper these principles are to be set forth as non-technically as possible; and a method for applying them to the purpose mentioned, requiring no intimate knowledge of the principles, will be described. The outcome of such application is a new chronoscope which, considering its simplicity, yields measurements of remarkable accuracy.

In considering the principles, an analogy from hydraulics is of service, namely the flow of water in a pipe. The strength of a current of water (gallons-per-second) multiplied by the time (seconds) during which the current flows gives the total quantity (gallons) which has passed through the pipe during the interval in question. Exactly the same relations hold in the case of electricity flowing in a circuit. The product of current strength (amperes) and time (seconds) gives the total quantity of electricity (coulombs) which has passed through the circuit during the time interval. This suggests the following possibility: let a current of known strength be sent, during a time interval to be determined, through a circuit in which there is a device for measuring the quantity of electricity—a device analogous to a water meter, to continue the comparison. The value of the time interval is obtainable from

[1] Physical Laboratory.

the relation above mentioned; expressed algebraically, $t = Q/I$, where t is the time interval, Q the quantity, and I the current.[1] The current might, of course, be measured by one of several methods; and the moving coil galvanometer, of the type possessed by most psychological laboratories, is an instrument which is applicable to the measurement of electrical quantity.

Any quantity of electricity of sufficient magnitude, when discharged through a galvanometer, produces a 'throw' or swing of the coil through a certain angle, whence the coil immediately begins its return towards its position of rest. The angle is a function of the total quantity which has passed. When the rate at which the quantity passes is constant, in other words, when the current is unchanging in strength during the interval in which it flows, the quantity of electricity is calculable from the observed angle of throw. This fact has been made the basis for the measurement of time intervals as long as five seconds and more with an accuracy of a few thousandths of a second.[2] When the time interval to be measured is short, say of the order of 1/30 of the time required by the coil to execute one complete vibration, the throw is, for practical purposes, directly proportional to the quantity, and therefore, with a given current strength, to the time interval. It thus becomes possible to construct a galvanometer scale to indicate time intervals directly, with any desired range up to half the time occupied by one complete vibration of the coil. For very short intervals, corresponding to those, for example, of reaction time, the scale will be one of nearly equal divisions.

The innovation, however, of constructing a direct-reading time scale for a galvanometer, is of small value *per se*. In addition, some method should be devised which will produce exact correspondence between the scale readings and the time

[1] We observe that the same relation holds for the current of water. If, for instance, we know that a current of I gallons-per-second is flowing in a pipe, during an unknown interval of t seconds, then a reading of the water meter, giving the number of gallons, Q, which passed during the interval, enables us to calculate the value of t by the formula above given. Thus, with a fixed strength of current, the reading of the water meter is proportional to the time interval during which the water was flowing.

[2] Klopsteg, *Physical Review*, 8, 195, 1916.

intervals represented by them. One method, as has already been intimated, would be to make such control depend upon measurements of electrical current and quantity. If there were no other way, however, the method would be inapplicable except in the hands of an experimenter trained in making such determinations. The simplification which renders the use of the galvanometer as a chronoscope practicable, even on the part of an experimenter unfamiliar with the principles, is a new fall apparatus which, with the aid of a simple switchboard for the electrical adjustment, enables one easily and accurately to control the galvanometer readings.

Description of Electrical Connections.—For the purpose of sending a constant current through the galvanometer during any short time interval, an arrangement, the most important features of which are shown in Fig. 1, is used. A lead storage

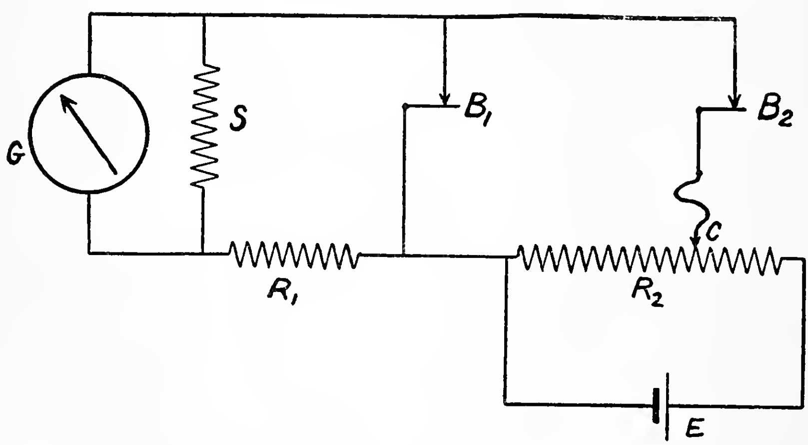


Fig. 1

cell is connected, as shown, to the terminals of a wire resistance R_2. By connecting one side of the galvanometer circuit to one of these terminals, and the other to a sliding contact C, which bears on the wire of R_2, we are enabled to impress any desired potential difference smaller than that of the cell upon the galvanometer, and thus cause a current of desired strength to pass through the instrument. In series with the galvanometer G is connected a high resistance R_1. To diminish the

likelihood of error in the measurement of an interval because of the fact that the current may not start exactly at the initial instant, nor be cut off exactly at the final instant of the interval, these instants are marked, preferably, by two breaks in the electrical connections, rather than by a make and a break. That is to say, the current should begin to flow through the galvanometer at the instant of breaking one connection, and cease at the instant of breaking a second connection. These are shown at B_1 and B_2 in the figure. The purpose of the high resistance R_1 is now evident: it renders the resistance of B_1 negligible in comparison with R_1 and the combination of G and S. This arrangement permits of no leak through the galvanometer before B_1 is opened. S is a shunt which produces critical damping of the galvanometer coil. In this condition the coil is, in the language of the engineer, 'dead beat,' and cannot pass the position of rest, or zero position, when returning from a throw.

Description of Complete Apparatus.—The complete chronoscope, arranged for the measurement of unknown intervals and for immediate determination of the accuracy of its indications, consists of the following three parts: galvanometer, fall apparatus and switchboard for electrical control, together with the necessary batteries. The component parts will be considered in detail.

Galvanometer.—In the writer's experimental setup, the galvanometer was one of the Leeds and Northrup type P instruments. Its time of throw, *i. e.*, the time required by the coil to make its excursion from the position of rest to its extreme displacement—which time is the same for all angles of throw—was adjusted to be about 2.5 seconds. With this time it is an easy matter to obtain the reading indicated at the extremity or turning point of the displacement. The coil was provided with a 'damping rectangle'[1] to assist in making the motion dead beat. The particular advantage of this degree of damping, in connection with the use of the galvanometer as a chronoscope, is the fact that the instrument

[1] Klopsteg, *Physical Review*, 3, 121, 1914; *Zeitschrift für Instrumentenkunde*, 34, 338, 1914.

is self setting to zero. A reading of the instrument is taken, let us say; by the time the observer has prepared for the following observation, the coil has returned to the zero position without oscillation, and is ready for the next reading. It is thus possible to obtain four or five observations a minute.

Two types of scale may be used. One which was found to be particularly convenient was translucent. Upon it the galvanometer throws were indicated by means of a vertical dark line extending through a spot of light. The spot, which was reflected from the mirror attached to the coil, was produced by an incandescent lamp attached to the scale, both scale and lamp having been affixed to the galvanometer. The other scale is the usual type, ruled on a white surface, and read by means of a telescope, the eyepiece of which is provided with a fine vertical crosshair. The advantage of the first type of scale is that of ease in reading, although the mean variation in a given number of measurements of the same interval may be somewhat greater than that resulting when the other type of scale is used. However, with this latter type, if the telescope is properly adjusted, a great number of readings can be taken without undue eye strain. Both scales were ruled for a range of 500 σ; on the translucent scale, each division represents 2 σ, while the other has one division per σ.

Fall Apparatus.—Bearing in mind that the initial and terminal instants of the interval to be measured are, at least in the control apparatus, to be marked by two breaks of electrical circuits, we observe this to be the problem in the design of the apparatus: to obtain exact coincidence between each of these two instants, respectively, and the corresponding break of a circuit. A fall apparatus, with a one-inch smooth iron or steel ball as the falling body, was chosen, for the reason that the distances of fall corresponding to certain time intervals can be accurately calculated, and that for the short intervals involved in these measurements, the air resistance to such a ball introduces no appreciable error. These circumstances—provided the coincidences above mentioned can be realized—render the fall apparatus independent of any other control, such as a tuning fork, for example. In Fig. 2 is shown the

apparatus which fulfilled the requirements with complete success. *E* is an electromagnet which, at the desired instant, upon the breaking of its energizing circuit, releases the iron ball. The latter, in falling, strikes the lever *A* which norm-

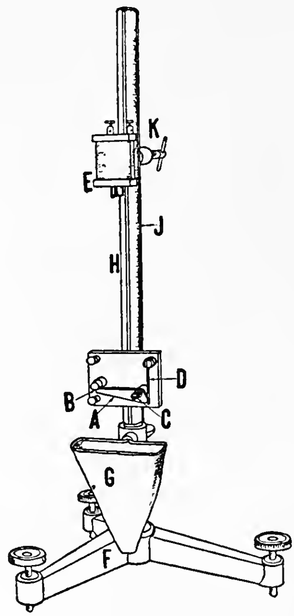


FIG. 2

ally makes contact at the point *B*; finally the ball is caught in the bag *G*. The release magnet, break switch and bag are all mounted on the rod *H* which, in turn, is supported by the tripod base *F*. The base is provided with levelling screws by means of which the axis of the magnet and middle point of the break lever can be brought into the same vertical line. The rod *H* carries the time scale *J* which indicates directly the time occupied in the descent of the ball from the magnet to the break lever. By means of the indicator *K* the position of the magnet is read on the scale *J*. The contact lever of the break switch pivots on the screw *C*. Electrical contact is assured by the fact that the leaf spring *D* bears upon the heel of the contact lever; on the other hand, after the falling ball has struck the lever, the latter is pushed out of the way and

held there by *D*, thus keeping the circuit open until the switch is intentionally closed. The range of the fall apparatus depends upon the length of the rod, but anything greater than 500 σ seems unnecessary, as will appear later; in fact 400 σ has been found sufficient for every purpose.

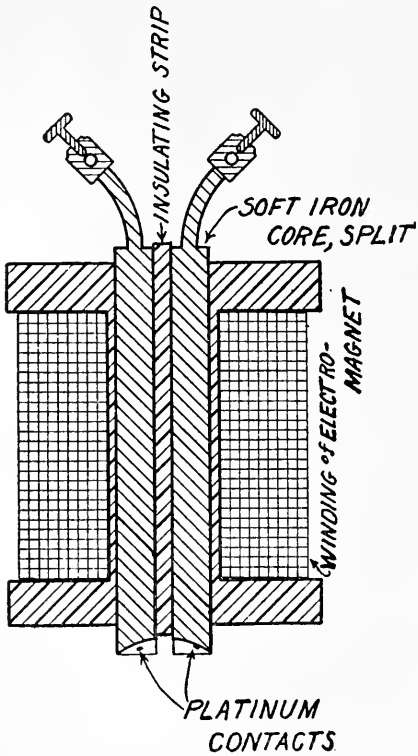


Fig. 3

In Fig. 3 is shown a cross-sectional diagram of the release magnet, which indicates how error in the time interval 'measured out' by the apparatus is obviated. As shown, the core of the magnet is split, with an insulating strip between the halves. Each of the halves is provided with a binding post. The winding is carefully insulated from the core. The latter is cupped at its lower end, so that it may hold the ball firmly. Three or four equally spaced platinum contacts are soldered into the cup, for good connection. It will be noticed

that when the ball is in place, and the halves of the core are connected in shunt with the galvanometer, the ball itself plays the rôle of B_1 in Fig. 1. All the current flows through the ball, not any through the galvanometer. Coincident with

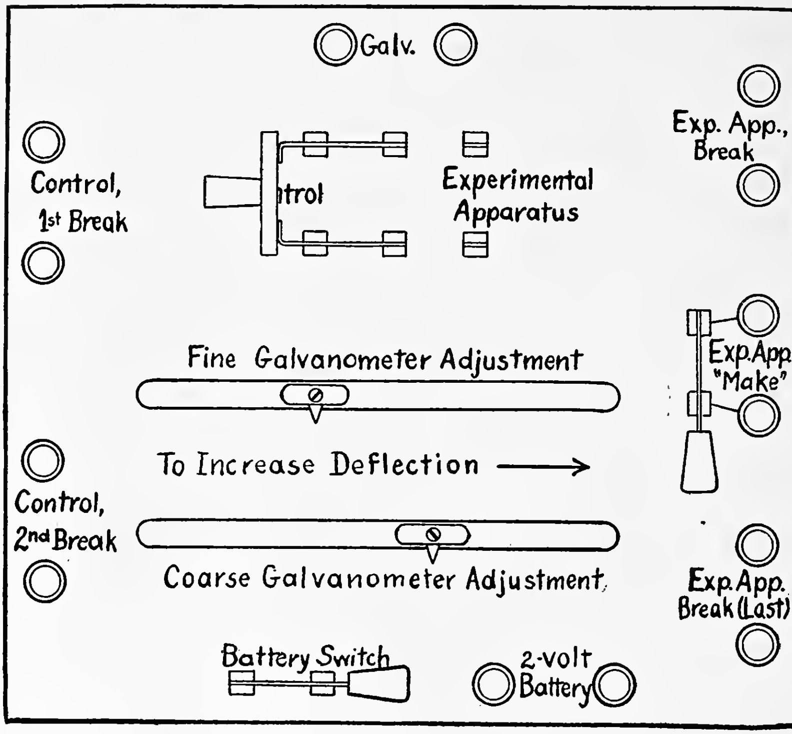


FIG. 4

the instant of release of the ball—and it should be emphasized that there is no other possibility besides coincidence—the current starts through the galvanometer, and continues until the instant at which the lever A is struck by the ball. The lever is represented by B_2 in Fig. 1. We are now able, by clamping the magnet at any desired position, as indicated on the scale J, to send a current through the galvanometer during the exact interval indicated.

Switchboard.—To make the manipulation of the chronoscope as simple as possible, the switchboard shown in Fig. 4

was designed.[1] All resistances indicated in Fig. 1 are mounted beneath the board. The entire board is so labeled that proper connections are easily made without possibility of error. With the double throw switch in the position shown, the control apparatus is connected to the chronoscope. Throwing this switch to the right connects the galvanometer and its necessary electrical attachments with, say, the reaction time apparatus, consisting presumably of a circuit for the stimulating mechanism, and of a reaction key circuit. It is best, in this apparatus, as in the control apparatus, to use two breaks. When this is not possible, a make may be used to effect the stimulation, followed by a break, made by the reaction movement.

Accuracy.—Although different galvanometers of the same type manifest individual differences, due principally to irregularities in their magnetic fields, still the performance of the instrument used by the author is quite typical of this particular kind. The galvanometer used in these measurements was not specially selected, but was taken because it was one of seventeen instruments not in use in the undergraduate laboratory. In taking the readings of the accompanying table, the fall apparatus was set for 300 σ, and the adjustment of the chronoscope made for zero constant error at this point. Having made this adjustment, the release magnet was set successively at 400, 350, 300, 200 and 100 σ, without further electrical adjustment, and ten readings were taken with the telescope and scale for each setting of the magnet. In the table the whole series of observations is given.

Particular attention should be called to the fact that for 100 σ, the distance the ball falls is but 4.903 cm. The observed values, considering that adjustment was made for zero constant error at 300 σ, are remarkably accurate, indicating that practically no error is introduced at the instants of starting and stopping the current by means of the apparatus. The results indicate that one is quite justified in assuming carefully made

[1] Improved designs of both release magnet and switchboard, involving the same principles as above described, have recently been worked out. The manufacture of the entire apparatus has been undertaken by the Leeds and Northrup Company of Philadelphia.

observations upon reaction times to give results which are safely within one σ of the true intervals. This will be especially true if the adjustment for zero constant error is made for an interval having somewhere near the mean value of the reaction intervals.

Fall Apparatus Set at—				
400	350	300	200	100
401.5	350.3	300	199.6	100.4
401.5	350	300	200	100
400.5	350	300.4	199.5	100
401.5	350	300	199.6	100
401	350	300	199.5	99.8
400	350.5	299.5	199.5	99.6
401.5	350	300	200	100
401	350	300	199.5	100.2
401	350	300	199.5	100
400.5	350	300	200	100
Av. 401	350.1	300	199.7	100
MV, σ . . . . *.42*	*.13*	*.10*	*.20*	*.12*

Summary

The present paper describes a simple form of chronoscope, especially suited to the measurement of intervals up to 500 σ, though adaptable to greater or smaller ranges. It consists of a galvanometer with a direct reading time scale, the indications of the scale corresponding to the intervals during which a steady current is permitted to flow through the instrument. The calibration of the scale depends upon the known relation between the total quantity of electricity which has passed through the galvanometer during a given time interval, and the length of this interval. As a means of adjustment and control of the scale readings, a new fall apparatus is described which accurately 'measures out' to the chronoscope any time interval within the range of the latter.

The advantages to be pointed out for this chronoscope with its fall apparatus are:

1. Great simplicity, combined with accuracy and ease of manipulation. Every part of the scale of the chronoscope is subject to immediate standardization, merely by throwing over a switch, and dropping a ball from a certain point, corre-

sponding to the point on the scale which is under test. The fall apparatus is constructed in such a way that no observable error can enter into the time interval which it determines. The probable accuracy of a single reading is easily within one σ of the true interval, and the mean variation in ten readings for the same interval is, for the most part, less than one half σ.

2. Absolute silence during operation. The only moving part is the galvanometer coil, suspended by means of an elastic metallic fiber.

3. Automatic return of the indicator to the zero reading, and rapidity of operation, making possible about four observations per minute.

4. Possibility of demonstrating measurement of short intervals to an audience. This merely requires a different arrangement of lamp and scale from that described in this paper—an arrangement such that the audience can easily follow the motion of the indicator, permitting each person to make his own observation.

5. The fall apparatus can be used apart from this chronoscope for obtaining an accurate check upon other devices.

Note added July, 1917.—Since the above paper was first prepared Dr. Herbert Woodrow and assistants have used the writer's apparatus for a series of about 6,000 observations upon sound reaction. The apparatus functioned perfectly throughout this work, during the course of which it was transported and set up in another laboratory. To Dr. Woodrow the writer wishes to acknowledge his gratitude for the time spent in giving him the psychologist's point of view in the design of the apparatus.

THE ESTIMATION OF DISTANCES BY SIGHT AND PASSIVE TOUCH: SOME INVESTIGATIONS INTO THE EVOLUTION OF THE SENSE OF TOUCH

BY ARTHUR B. FITT, M.A., Ph.D.

Head of Education Department, Rhodes University College, Grahamstown, South Africa

Contents

I. (*a*) Introduction

This investigation is a direct continuation of the author's article 'Grössenauffassung durch das Auge und den ruhenden Tastsinn,' published in *Archiv für die gesamte Psychologie*, Vol. XXXII., p. 420 ff.

The first part of this article will be devoted to an attempt at a biological explanation of the general law from the point of view of adaptation. The second part will present the results of fresh experiments conducted with children of various ages, further verifying the general law and its biological significance.

So as to establish continuity between the earlier investigation and this, a brief résumé of the main points will be given.

(*b*) METHODS OF EXPERIMENT, APPARATUS, FORMULÆ, ETC.

The general problem is: How accurately do we estimate distances (above the space-threshold) presented to passive touch by means of two, simultaneously given, unmoved points?

The tactile distance was compared with the visual in each case, the presentation of the two impressions being successive. The method used was the Constant-method, with full rows of the variable stimulus, the distances following each other not consecutively, but as drawn by chance. As a rule 5 rows of the variable stimulus formed the basis of each calculation for the equivalent-value.

To calculate this equivalent-value the following Spearman[1]-Wirth[2] formulæ were used:

$$\text{Upper Threshold, } r_0 = E_0 - i\Sigma g + \frac{i}{2},$$

$$\text{Lower Threshold, } r_u = E_u + i\Sigma k - \frac{i}{2},$$

wherein Σg is the sum of all the 'greater' and Σk of all 'smaller'-judgments, in their relative frequency, E_0 and E_u are the first extremes beyond which no more 'greater' or 'smaller'-judgments are given, and i is the interval between consecutive distances in the variable stimulus.

$$\text{The equivalent-value } A = \frac{r_0 + r_u}{2}.$$

For calculation of the spread or distribution of the decisions ('Streuungsmasse'), Wirth's formulæ[3] were used. M_0 is the value of the spread or 'Streuung' with reference to the upper threshold, M_u with reference to the lower threshold, and M is the value of the total 'Streuung.'

$$\frac{1}{i^2}(M_0)^2 = 2[(p-1)g_1 + (p-2)g_2 \cdots + 1\cdot g_{p-1} + \tfrac{1}{8}] - (\Sigma g - \tfrac{1}{2})^2,$$

[1] *Brit. Jour. of Psy.*, Vol. II., 3, 1908, p. 227 ff.

[2] *Psychophysik*, 1912, pp. 188 and 189.

[3] Wirth, *Psychophysik*, 1912, p. 192.

$$\frac{1}{i^2}(M_u)^2 = 2[(q - 1)k_1 + (q - 2)k_2 \cdots + 1 \cdot k_{q-1} + \tfrac{1}{8}] - (\Sigma k - \tfrac{1}{2})^2,$$

$$M = \sqrt{\tfrac{1}{2}(M_0^2 + M_u^2 + 2S^2)},$$

in which $2S = r_0 - r_u$.

The visual distance was presented by means of two black dots on a white card which stood perpendicularly at normal seeing-distance from the observer. In the first investigation and in the experiments with the German boys the tactile stimulus was given by means of an Ebbinghaus æsthesiometer (applied electrically), but by means of a Spearman æsthesiometer (applied by hand)[1] in the experiments with the Grahamstown boys embodied in this investigation. Whilst the tactile stimulus was being given, the card with the dots was covered. Throughout each series of experiments the æsthesiometer and the stimulated portion of the skin were hidden from view by a screen. The time-order, which was regulated in both investigations by a metronome, was as follows:

Interval between signal ('Bitte,' 'ready,' etc.) and the first stimulus, and between the first and second stimulus: each one second. Duration of each stimulus: as a rule 2 seconds.

The details of the stimulation were unknown, except that the observer knew which part of the skin was to be stimulated and in which direction (across or along).

The order of succession of the two stimuli which was found most successful with the adults, viz., visual stimulus as normal stimulus followed by tactile stimulus as variable, was adopted throughout with the boys.

The total length of the sittings varied from about 5 minutes with the 6-year-olds to about 20 minutes with the 14-year-olds. Even these were broken up into smaller periods, as a rule two or more boys relieving one another in turn.

[1] Careful manual stimulation is quite adequate for ordinary purposes, a view recently supported by Carnes & Shearer, in *Amer. Jour. of Psy.*, Vol. XXVII., 3, July, 1916, p. 417 ff.

II. (*a*) The General Law of the Relation between the Space-threshold and the Estimation of Two-point Distances

Those parts of the skin which have very small space-thresholds overestimate two-point distances, the overestimation gradually decreasing with increase of the threshold until the indifference-point is reached where the estimation is correct. Increase of the threshold beyond this point brings underestimation, and the greater the increase the greater the underestimation.

In order to represent some of the original details, the average results for the adults and those for one of the individuals *M* will be reproduced here.

The results given are for the following parts of the skin:

1. Volar side of the first phalange of the forefinger-middle-lengthwise-right hand (short 'Forefinger').

2. Volar side of the second phalange of the second finger-middle-lengthwise-right hand (short 'Second finger').

3. Mounds or cushions of the hand-across-right hand (short 'Cushions of hand').

4. Back of hand-across-about 35 mm. from the knuckles-right (short 'Back of hand').

5. Volar side of the forearm—between middle and wrist—lengthwise towards thumb-side—right (short 'Forearm').

6. Back of neck—little to right of center—lengthwise (short 'Back of neck').

In the experiments with the boys, the parts of the skin were 1, 3, 4 and 5 of the above, making allowance in 4 for the difference in the size of the hand.

The symbols M_0, M_u, M, E_0, E_u, r_0 and r_u in the tables are as explained in the formulæ above. N is the length of the visual distance which was given first as constant stimulus and A the calculated value of the tactile distance which was judged equal to it. The amount of error involved in the tactile estimation, calculated to a percentage, is indicated briefly as 'error.' An overestimation is indicated by + and an underestimation by −. All measurements are given in millimeters.

Table I

Observer M

Skin-locality	Space-threshold	N	A	Error Per Cent.	M_o	M_u	M	E_o	r_o	E_u	r_u
1. Forefinger	2.65	18	16.8	+ 6.7	4.08	3.74	4.12	20	18.1	14	15.5
2. Second finger	7.5	9	8.8	+ 2.2	2.29	3.93	3.42	11	10	6	7.6
3. Cushions of hand	11.4	32	32.3	− 1	2.93	6.77	5.40	35	33.7	28	30.9
4. Back of hand	18.2	35	37.9	− 8.3	4.66	6.62	5.81	41	38.9	34	36.9
5. Forearm	34.6	58	71.4	−23.1	7.63	5.74	7.31	80	74.2	64	68.6
6. Back of neck	48.25	38	60.1	−58.2	10.39	7	9.18	67	62.5	55	57.7

Table II

Averages for Adults

Skin-locality	Space-threshold	Error Per Cent.	M_o	M_u	M
1. Forefinger	2.23	+25.2	4.11	3.70	4.24
2. Cushions of hand	6.83	+ 1.8	4.62	4.87	5.02
3. Back of hand	16.83	−10.3	8.73	8.40	8.82
4. Forearm	28.6	−58.8	15.40	21.75	19.30

(*b*) The Relation between the Accuracy of the Estimation and the Ability to Visualize the Pressure-points

The main factor in the explanation of the general law was the ability for 'passive visualization' of the points of the skin stimulated.[1] This varies inversely to the size of the threshold. As an immediate experience it appears in the different degrees of 'liveliness' of the tactual impression. Thus the variable law of estimation is really a statement of the different degrees of coördination between sight and pressure corresponding to the different parts of the skin-surface. The parts with the closest coördination exaggerate the two-point distance and those with the least coördination underestimate.

(*c*) A Synthesis between Vierordt's Law of the Relation of the Space-threshold to the Mobility of the Part and the Law of Tactual Estimation of Distances

Vierordt's law[2] states that the more mobile the part of the body is, the smaller is the threshold. This means in terms of my law: The more mobile the part, *i. e.*, the more the

[1] Fitt, 'Grössenauffassung, etc.,' p. 441 ff. and p. 445.

[2] Vierordt, in *Zeitschrift für Biologie*, VI., 1870, pp. 53–72.

part is and can be used, the higher does the estimation of distances rise towards or in the zone of overestimation. Thus those parts which are of most importance in measuring distances tend towards exaggeration of the magnitudes.

(d) Biological Significance of the Law. Different Degrees of Adaptation of the Various Parts of the Skin According to Mobility or Use

Since the sense of touch cannot alone give the same fulness of knowledge of the external world as, *e. g.*, sight, it is natural to expect some adaptation bringing compensation with it. The means of adaptation is the different degrees of pressure-sight coördination. The adaptive compensation is the ability to exaggerate magnitudes, thus to get a very vivid impression from external objects, this ability increasing with the need of adaptation to this mode of measuring.

The tip of the tongue, which has the smallest threshold, has undoubtedly the most vital decisions to make concerning the size, shape, etc., of very minute objects received into the mouth. How usefu it is biologically that a tremendous overestimation should take place! Then again how vital are many of the decisions to be made in the dark through the knowledge gained from the finger-tips!

Thus from the point of view of evolution, through usage in the life of the race the various parts of the skin have become differentiated in regard to accuracy of estimation from what was once almost certainly a general underestimation to the state of the present variable law.

III. (a) The Experiments with Boys; the Development of the Tactual Estimation from the Age of 6 Years to the Adult Stage

The experiments with the Grahamstown boys (white) were conducted in the years 1914 and 1915 and those with the German boys in the year 1913 in the Institute for Experimental Pedagogy in the University of Leipzig. Of the former there were 5 boys of 6 years, 4 of 8 years, 4 of 10 years, 5 of 12 years, and 4 of 14 years (total 22) and of the latter 4 of

10 years and 4 of 14 years. As complete results as possible were obtained from the 6-, 10- and 14-year-old Grahamstown boys, but from the 8- and 12-year-olds only some test results to verify or modify (if necessary) the general tendency in the development. The results from the German boys are complete, but with only 2 stages in the development, alone they are of little or no use. They prove particularly interesting, however, when compared with the other more complete series.

The tables which now follow present the complete results for the Grahamstown 6-, 10- and 14-year-olds, first the individual results, then the averages to be compared with the averages for adults. It must be remarked here that the boys were in every case chosen merely at random, so that one should hope to expect a fairly average result for each age, without of course any certainty of the most general results. It has cost much labor and time to obtain this material, and further there is no indication that a greater number of observers was necessary.

TABLE III

6-YEAR-OLD BOYS

Skin-locality	Sp.-Thr.	N	A	Error %	M_o	M_u	M	E_o	r_o	E_u	r_u	Obs.
1. Forefinger.......		10	8.6	+14	5.00	3.16	4.18	11	8.7	7	8.5	B.
		10	9.7	+ 3	2.25	3.16	2.75	11	9.9	8	9.5	M.
		10	7.9	+21	9.02	5.35	7.42	12	7.9	6	7.9	R.
		10	8	+20	5.46	4.54	5.03	10	8.1	6	7.9	S.
		10	7.5	+25	3.16	3.16	3.32	10	8.5	5	6.5	Wb.
2. Cushions of hand.		15	19.66	−31.1	3.94	2.83	3.44	22	19.8	18	19.5	B.
		15	16.66	−11.1	2.63	4.25	3.53	18	16.8	14	16.5	M.
		15	15.83	− 6.2	1.25	2.83	2.21	17	16.2	14	15.5	R.
		15	16.5	−10	1.71	3.68	2.89	18	16.8	14	16.2	S.
		15	14.66	+ 2.2	4.25	1.25	3.62	19	16.5	12	12.8	Wb.
3. Back of hand....		15	20.2	−34.5	6.13	3.30	4.91	24	20.2	19	20.2	B.
		15	21.33	−42.2	2.83	5.32	4.26	23	21.5	18	21.2	M.
		15	20.33	−35.5	0	4.92	3.49	21	20.5	17	20.2	R.
		15	19	−26.7								S.
		15	19.66	−31.1	4.86	2.83	4.52	24	21.8	16	17.5	Wb.
4. Forearm........		20	40	−83.3								B.
		15	25									
		20	32	−66.66								M.
		15	26									
		20	68	−240								R.
		20	38	−115								S.
		15	36									
		15	26	−61.7								Wb.
		20	30									

Table IV

10-Year-old Boys

Skin-locality	Sp.-thr.	N	A	Error Per Cent.	M_o	M_u	M	E_o	r_o	E_u	r_u	Obs.
1. Forefinger......	1.38	10	11.4	− 14	4.40	3.16	3.94	14	12.3	9	10.5	H.
	1.5	10	9.1	+ 9	2.65	5.56	4.43	11	9.9	6	8.3	Th.
	2.25	10	7.5	+ 25	1.85	1.85	1.95	9	8.1	6	6.9	Tr.
		10	13.5	− 35	5.66	7.75	6.86	17	14.5	9	12.5	Ws.
2. Cushions of hand	8.25	15	20.7	− 38	3.32	9.70	7.57	24	22.9	14	18.5	H.
	4	15	20.3	− 35.3	6.71	5.38	6.16	24	21.3	17	19.3	Th.
	6	15	17.7	− 18	4.40	9.35	7.33	20	18.3	13	17.1	Tr.
		15	27.6	− 84	6.71	5.16	6.22	32	29.3	24	25.9	Ws.
3. Back of hand...	14	15	29.7	− 98	8.83	9.33	9.26	35	31.5	24	27.9	H.
	13	15	23.5	− 56.7	6.55	8.91	7.91	27	24.7	19	22.3	Th.
	12.5	22	27.8	− 26.4	5.48	9.24	7.63	31	28.5	23	27.1	Tr.
		15	24.5	− 63.3	8.21	10.29	9.38	29	25.7	19	23.3	Ws.
4. Forearm.......	29	10	41.66	−316.7	20.98	10.88	17.53	58	47	32	36.3	H.
		10	45	−350								Th.
	33.5	22	61.33	−178.8	22.25	20.59	21.44	74	61.7	50	61	Tr.
		22	67.33	−206	25.14	24.71	25.19	82	71	52	63.7	Ws.

Table V

14-Year-old Boys

Skin-locality	Sp.-thr.	N	A	Error Per Cent.	M_o	M_u	M	E_o	r_o	E_u	r_u	Obs.
1. Forefinger......	1.92	14	14.75	− 5.3	8.20	8.49	8.44	19	16	10	13.5	Bo.
	1.75	14	16	− 14.3	5.68	10.40	11.94	19	17	11	15	Br.
	1.38	14	12.92	+ 8.4	5.05	3.74	4.66	16	14.3	10	11.5	P.
	1.5	14	9.6	+ 31.4	8.40	1.90	6.32	15	11.3	7	7.9	Ts.
2. Cushions of hand	6	15	20.63	− 37.5	4.92	7.05	6.23	24	22	16	19.25	Bo.
	6.5	32	48.1	− 50.3	6.08	8.25	7.79	53	50.7	42	45.5	Br.
	6.75	32	31.4	+ 1.9	6.40	6.00	6.49	36	33.3	27	29.5	P.
	6.75	32	33.5	− 4.7	7.28	9.09	8.43	38	35.3	28	31.7	Ts.
3. Back of hand...	14.7	28	33.75	− 20.5	5.58	10.08	8.39	38	35.75	27	31.75	Bo.
	15.83	15	45.5	−203.3	9.06	14.56	12.29	51	47.5	36	43.5	Br.
	16	28	28	0	9.00	7.20	8.42	34	30.1	23	25.9	P.
	16.25	28	36.7	− 31.1	6.85	7.26	7.24	41	38.3	32	35.1	Ts.
4. Forearm.......	23.7	20	52.6	−163	15.98	22.27	19.45	60	54.2	42	51	Bo.
	23.7	20	71.8	−259	14.95	38.91	29.57	80	74.2	52	69.4	Br.
	23.5	40	60.4	− 51	20.25	23.28	22.22	72	64.6	48	56.2	P.
	29	20	44.4	−122	18.33	25.94	22.93	56	49	30	39.8	Ts.

On glancing at Table VI, one is struck immediately by the fact that the general law holds definitely for each age, *i. e.*, each part of the skin has its own law of estimation and the relative order of the various points is well maintained throughout.

Table VI

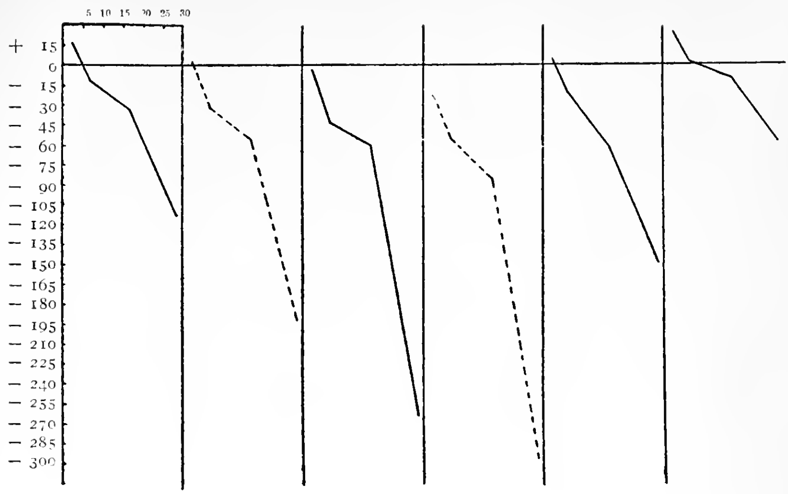
Averages for Boys and Adults

Skin-locality	Space-threshold	Error Per Cent.	M_o	M_u	M	Age
1. Forefinger		+ 16.6	4.98	3.87	4.54	6 years
2. Cushions of hand		− 11.2	2.76	2.97	3.14	
3. Back of hand		− 34.0	3.46	4.09	4.30	
4. Forearm		−113.3				
1. Forefinger	1.7	− 3.75	3.64	4.58	4.30	10 years
2. Cushions of hand	6.1	− 43.8	5.29	7.40	6.82	
3. Back of hand	13.2	− 61.1	7.27	9.44	8.55	
4. Forearm	31.25	−262.9	22.79	18.73	21.39	
1. Forefinger	1.64	+ 5	7.51	6.13	7.84	14 years
2. Cushions of hand	6.5	− 22.7	6.17	7.60	7.24	
3. Back of hand	15.7	− 63.7	7.62	9.78	9.09	
4. Forearm	25	−148.8	17.35	27.60	23.54	
1. Forefinger	2.23	+ 25.2	4.11	3.70	4.24	Adults
2. Cushions of hand	6.83	+ 1.8	4.62	4.87	5.02	
3. Back of hand	16.83	− 10.3	8.73	8.40	8.82	
4. Forearm	28.6	− 58.8	15.40	21.75	19.30	

The first point of difference is that the estimation-curves are differently situated round the indifference-point. It is remarkable that the 6-year-olds (not long entered school) estimate almost as the adults do, although in terms of the general evolutionary interpretation outlined above[1] still somewhat below their level. The whole curve is much longer and much more in the minus-zone than the adult curve. See curves 1, 3, 5 and 6 which correspond to Table VI.[2] From the 6-year-olds to the 12-year-olds, instead of a rise, there is a fall in the estimation curve. Then it rises in the case of the 14-year-olds, although not up to the level of the 6-year-olds, for the 14-year curve is slightly longer and each point lies below the corresponding point in the 6-year curve. The further rise to the adult curve is considerable. So regularly do the four points in the curves rise and fall that the position of any one point, once determined, roughly fixes the position of the whole curve. The investigator, on throwing the results together and finding that the development was a broken one, wished to pursue the investigation further, to the extent of

[1] P. 269.

[2] P. 273.



Curve 1, 6 yrs. Curve 2, 8 yrs. Curve 3, 10 yrs. Curve 4, 12 yrs. Curve 5, 14 yrs. Curve 6, Adults.

The abscissæ represent the space-threshold values and the ordinates the error per cent. + indicates exaggeration and — underestimation.

testing the intermediate periods at the 8- and 12-year-old stages. Values were found for one part of the skin, 'cushions of hand,' and from these proportionate curves (curves 2 and 4) were drawn. There is a high degree of probability that these curves represent the approximate estimation at these two stages. The estimation values for the part ('cushions') are:

8-Yr.-Olds. $N = 15$				12-Yr.-Olds. $N = 15$				
A-values				*A*-values				
Obs. *A*	*B*	*C*	*D*	Obs. *A*	*B*	*C*	*D*	*E*
20.5	16.5	22	20.5	24	21.5	21	21.5	29
Average 19.9 = Error per cent. — 32.7				Average 23.4 = Error per cent. — 56				

The test shows that the estimation for this part of the skin with the 8-year-olds lies in an intermediate position between the corresponding ones with the 6- and 10-year-olds and on the falling grade. The 12-year-old value lies below the 10-year one, so that the rise to the upgrade has not yet commenced. A glance at the curves would suggest some-

where about 13 years as the turning-point for the change upward. The rates of falling and rising are very different. The fall for each period of 2 years from the 6- to the 12-year-olds is slight and practically equal, but the rise is rapid, and particularly so from the changing point to the 14-year stage.

A study of the averages for the various ages has thus shown a regular development. This is demonstrated just as clearly when one compares the individual cases. That the 6-year-olds certainly show a more advanced development than the adults is demonstrated by the fact that the most advanced 6-year-old, Wb., presents a curve practically the same as the average adult curve.

Wb.	+25	+2.2	−31.1	−61.7
Adults	+25.2	+1.8	−10.3	−58.8

Wb.'s curve does not stand nearly as high, however, as several of the most advanced individual adult curves, *e. g.*, Kr., Pc., etc.

The devolution from the 6- to 10-year-olds is demonstrated in the fact that two of the latter show underestimation for all parts of the skin (and even a great underestimation for the first part, forefinger), this occurring in none of the 6-year cases. One sees it further in the great degree of underestimation for the forearm in all 10-year cases as compared with the much smaller minus-values for the same part in the 6-year cases.

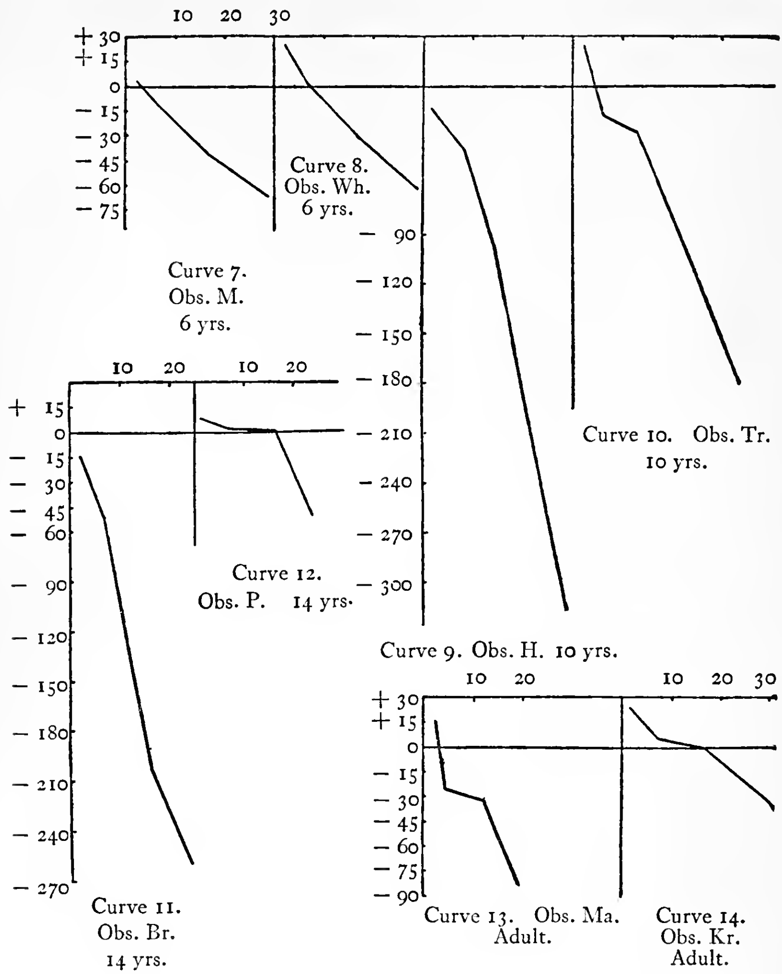
One sees the distinct advance from the 10-year to 14-year-olds in the fact that two of the latter, P. and Ts., present considerably more advanced curves than the most advanced of the former, Tr.

14 yrs.	P.	+ 8.4	+1.9	0	− 51
	Ts.	+31.4	−4.7	−31.1	−122
10 yrs.	Tr.	+25	−18	−26.4	−178.8

The least advanced cases in these two groups, Ws. and Br., are practically the same.

Although the 14-year average is far behind the adult average, and further two of the 14-year individuals, Br. and Bo., show a minus-value for the forefinger (and correspondingly

large minus-values for the other parts) which occurs in no adult case, still it is interesting to note the advanced 14-year case



of P. which is comparable to some of the advanced adult cases, *e. g.*, Kr. and Kn.

14 yrs. P.	+8.4	+1.9	0	−51
Adults Kr.	+22.8	+4.7	−0.6	−39.8
Adults Kn.	+37.2	−5.3	−1.1	−22.9

Possibly P. is already advanced as far as he will go in the scale of estimation.

Clearly the average 14-year case has not yet reached the level of the average 6-year-olds. The two backward 14-year

cases of Bo. and Br. lie apparently at about the point of minimal development, *i. e.*, at the turning-point from devolution to advance or very little above it.

In order to illustrate this development through the individual results, curves 7–14 are given, representing the least and the most advanced case for each period (as far as it is possible to choose so definitely).

The Precision-Values

An explanatory remark concerning the M_0-, M_u- and M-values for the 6-year-olds is necessary here. They are very small and this is possibly partly accounted for by the fact that as a rule not as many rows of the variable stimulus for each A-value were given as with the other observers. This was found necessary in order to keep the interest and hence effort in the experiments from flagging. Still it was not considered detrimental to the general results, for the experimenter soon saw from the smallness of the decision-times and the general readiness and ease in the decisions that the precision-values were particularly high. This was not so in the experiments with the fourth group (forearm) where the judgments came with such difficulty that it was considered unwise to attempt full rows of the variable stimulus. Hence the results for this group were obtained from several judgments for distances lying differently about the A-point, without however sufficient data to determine r- and E-values. The above procedure of giving fewer rows of the touch-stimulus had to be adopted occasionally in the experiments with the older boys, but there is no indication of smaller M-values than would otherwise have been obtained. Thus from every indication one can safely state that the precision-values for the 6-year-olds (except for the forearm) are comparatively high, *i. e.*, that the judgments are relatively certain.

There is a clear rise in all the M-values from the 10- to the 14-year-olds, *i. e.*, a fall in the precision. Just as clear too is the rise in the precision from the 14-year-olds to the adults.

Thus there is most probably, if not a fall, from the precision-values of the 6-year-olds to those of the 10-year-olds, at least

no rise. Then there is a fall in the 10–14 years period with a rise to the adult stage. We see here an analogous development to that of the estimation, with a devolution to a certain point, and then a rise to the adult stage. The point where this change to the rise occurs is, however, different in each case. Whereas the estimation-curves rise at about 13 years, the precision-values rise after 14 years. The relation of these two facts will be discussed later. This development of the precision is shown in the following table. The value given for each age is the average struck from the h-values (precision-values) for the 4 parts of the skin. This is quite allowable, since the h-values for the 4 parts lie in a definite relation to one another.

TABLE VII

$$h = \frac{I}{M\sqrt{2}}.$$

	10 Years	14 Years	Adults
h-values	.07	.06	.08
M-values	10.27	11.93	9.35

The Space-Threshold Values

It is generally assumed that the space-thresholds are smaller with children than with adults, and increase gradually with the age. Our previous investigation established the fact that the actual individual differences in size of the thresholds do not influence the general law of estimation, hence it is practically immaterial here to have the values of the thresholds for the various groups of boys. However, as far as possible, they have been taken, and given in the above tables. The usual procedure for children was adopted here, the procedure of making the end-points in the ascending and descending series 'two-point' and 'one-point' perceptions respectively, and not 'two-point' and 'no longer distinctly two points' as with the adults. This naturally gives slightly smaller results than would be obtained on the adult procedure, and thus the thresholds of the adults and boys are not really comparable. Even with this simpler procedure, much difficulty was found in obtaining satisfactory results, and with the 6-year-olds the

attempt to obtain anything consistent had to be abandoned. Thus the threshold-values for the boys are incomparable with those of the adults both on account of the difference in methodical procedure and also because of their greater uncertainty. For these reasons all curves for averages have as their abscissæ the same threshold-values, those of the adults. As has been shown, this does not affect the general law.

(*b*) The Significance of this Development

These investigations with the boys give considerable support to the general theory of the evolution of the estimation as outlined above. According to this, there appears from the sixth to about the thirteenth year a diminution in the liveliness of the touch-impressions, and at the same time a breaking-down of the coördination between touch and sight, followed by a rapid building-up of this coördination. From the point of view of adaptation we see a fall from the 6-year curve to the 10-year one, the fall continuing to about the thirteenth year, when a rapid regaining sets in, the adult stage showing the highest degree of adaptation. The 6-year curve shows a fair portion above the indifference line, *i. e.*, a relatively close approximation to correct estimation by many parts, with exaggeration by the most sensitive surfaces. The adaptation at 10 years is poor, for all parts show underestimation, which is very considerable with the least sensitive surfaces.

One can hardly resist the suggestion that the breaking-down of the touch-sight coördination is due to the nature of the mental activities of school-life; to the sudden change from the incessant handling and seeing of infancy to the comparative cessation of this and the substitution of a greater activity in abstraction. That this breaking-down does not continue beyond somewhere about the thirteenth year suggests the influence of a new factor. The most likely one is the pubertal change which in many activities is a reinforcing and accelerating factor. It seems natural that the arrest in coördination should not be allowed to continue beyond this vitally important period, and that a rapid building-up should commence. The

thirteenth year is most probably with South African boys on the average the commencing year for the pubertal changes. No very exact statistics for this fact are available, but one may almost assume it from the fact that the average year for English boys is about the fourteenth and that a warmer climate as a general rule means earlier pubescence. The main point for this investigation is the fact that the 14-year-old boys who acted as observers were found to be in every respect some distance on in the pubertal changes and not merely commencing.

From the point of view of sense-development the breaking-down of the coördination, caused almost certainly by school-entrance, is a criticism of the tendencies of modern school-life. It also strongly supports some of the newer types of educational institutions, *e. g.*, the Montessori and manual-training schools.

A breaking-down in the touch-sight coördination would tend to bring about a fall in the precision in the tactual estimation of distances. Were this the only factor involved in the precision one would expect to find exactly parallel developments of the two. Of course there are many factors, one of them being of outstanding importance, the increase or decrease of the general capacity for concentration apart from any specific reference to the nature of the judgment itself. The numerous investigations on the pubertal change incline one to the view that the early stages of puberty are unfavorable to a high degree of precision in the judgment-process. With this fact in mind, one can see considerable parallelism between the development of the estimation and the degree of precision. In the latter there is strong indication of a fall to the fourteenth year (and beyond?), in the former a decided fall to about the thirteenth year with then a rise in each case to the adult stage. The change from devolution to advance has been attributed in each case to the same factor, the pubertal change.

The author wishes to draw special attention to the comparatively high degree of precision with the 6-year-olds. This he attributes to the fact that the boys, living very much in the world of sense, give their decisions frankly and freely, without the scruples which the higher thought-process is liable to

bring. As well, of course, the close touch-sight coördination plays its part.

Experiments with German Boys

For purposes of comparison, and if possible, verification of the generalizations above arrived at, the experiments conducted with the German boys will be given here. The results of Table X. (averages for 10- and 14-year-olds compared with adult averages) are represented in Curves 15, 16 and 17, Curve 17 being Curve 6 reproduced for ease of comparison.

TABLE VIII

10-YEAR-OLD BOYS

Skin-locality	Sp.-thr.	N	A	Error Per Cent.	M_o	M_u	M	E_o	r_o	E_u	r_u	Obs.
1. Forefinger....	1.9	14	9.33	+ 33.3	6.83	3.24	5.72	14	11.33	6	7.33	S.
	2.8	14	12.83	+ 8.4	4.00	5.44	5.06	16	14.5	9	11.17	B.
	2.1	14	13.16	+ 6.0	4.41	5.12	4.92	16	14.33	10	12.0	J.
	2.2	14	11.58	+ 17.3	6.23	8.83	7.64	14	11.67	8	11.5	F.
2. Cushions of hand......	4.9	32	30.92	+ 3.4	5.79	7.16	6.79	35	32.83	26	29.0	S.
	7.8	32	32.42	− 1.3	6.54	8.84	8.11	38	35.67	26	29.17	B.
	3.2	32	33.00	− 3.1	7.96	6.24	8.05	40	36.7	27	29.3	J.
	7.8	32	33.40	− 4.4	6.16	7.31	6.76	36	33.5	30	33.3	F.
3. Back of hand.	10.7	45	47.58	− 5.7	9.70	11.92	11.25	54	50.5	40	44.67	S.
	18.2	45	49.75	− 10.6	13.78	13.61	14.0	57	52.5	42	47.0	B.
	21.0	38	41.50	− 9.2	11.15	10.40	11.0	48	44.0	35	39.0	J.
	14.0	45	47.50	− 5.6	10.30	10.30	10.30	51	47.5	44	47.5	F.
4. Forearm.....	17.4	40	60.67	− 51.7	32.32	51.80	43.6	78	67	34	54.33	S.
	35.0	40	53.60	− 34.0	18.82	27.14	23.9	66	58.6	38	48.6	B.
	27.3	40	46.00	− 15.0	33.58	11.63	26.3	68	53.8	34	38.2	J.
	23.0	40	61.00	− 52.5	22.80	29.53	26.4	70	61.0	48	06.1	F.

A first glance at these results certainly does not convince one of their similarity to the previous ones as regards development. One is struck first by the shortness of the curves for the 10- and 14-year-olds as compared with the adult curve. Then one finds no such thing as a great rise from the 10-year to the 14-year curve, but rather a suggestion of a slight fall. On looking more closely however, and bearing in mind the different age of the pubertal changes and the very different conditions of life, in school and out of school, one begins to see similarities. Both 10- and 14-year-olds exaggerate with

Table IX

14-Year-old Boys

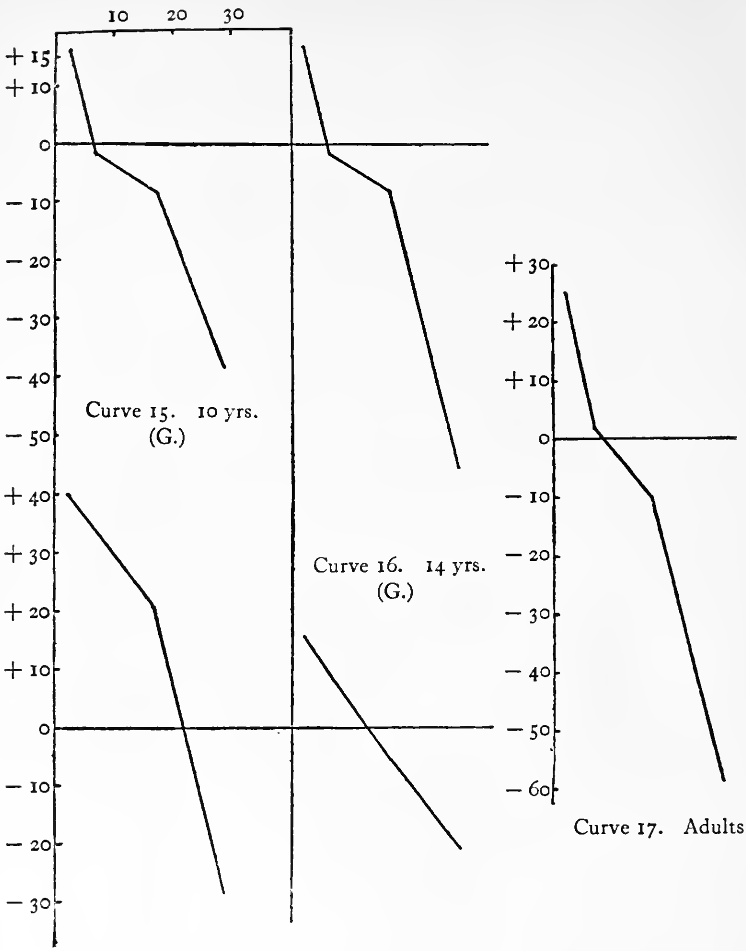
Skin-locality	Sp.-thr.	N	A	Error Per Cent.	M	M_n	M	E_o	r_o	E_u	r_u	Obs.
1. Forefinger....	2.1	14	8.88	+ 36.6	4.79	4.41	4.77	12	10.17	6	7.7	Lö.
	2.0	14	12.17	+ 13.1	6.41	4.92	5.84	16	13.33	9	11.0	Lu.
	1.9	14	13.0	+ 7.1	4.40	6.08	5.46	16	14.3	9	11.7	Lr.
	2.0	14	12.6	+ 10.0	5.74	5.66	5.83	16	13.7	9	11.5	K.
2. Cushions of hand......	6.8	32	29.08	+ 9.1	10.21	6.90	8.92	35	31.0	25	27.17	Lö.
	9.0	32	33.92	− 6.0	4.52	10.34	8.17	37	35.67	28	32.17	Lu.
	6.8	32	34.8	− 8.8	6.71	5.74	6.42	39	36.3	31	33.3	Lr.
	7.2	32	32.2	− 0.6	5.42	5.28	6.2	37	35.3	27	29.1	K.
3. Back of hand.	13.9	45	46.5	− 3.3	12.01	11.33	12.2	54	50.0	39	43.0	Lö.
	20.0	45	48.5	− 7.8	9.23	21.18	16.5	54	50.67	38	46.33	Lu.
	16.3	45	52.1	− 15.8	12.01	10.07	11.3	59	54.1	46	50.1	Lr.
	11.9	45	47.8	− 6.2	7.87	15.20	12.65	54	51.5	38	44.1	K.
4. Forearm.....	31.4	40	47.0	− 17.5	28.01	13.09	22.3	62	51.33	38	42.67	Lö.
	30.0	40	43.3	− 8.3	13.83	8.00	12.7	52	49.1	36	37.5	Lu.
	34.3	40	61.8	− 54.5	21.41	24.68	23.4	72	65.8	48	57.8	Lr.
	21.7	20	48.4	−142.0	18.83	26.53	23.65	60	53.8	34	43.0	K.

Table X

Averages

Skin-locality	Sp.-thr.	Error Per Cent.	M_o	M_u	M	
1. Forefinger..................	2.25	+16.25	5.37	5.66	5.84	10 years.
2. Cushions of hand............	5.93	− 1.35	6.61	7.39	7.43	
3. Back of hand................	15.98	− 7.8	11.23	11.56	11.64	
4. Forearm.....................	25.68	−38.3	26.88	30.03	30.05	
1. Forefinger..................	2.0	+16.7	5.34	5.27	5.48	14 years.
2. Cushions of hand............	7.45	− 1.6	6.72	7.07	7.14	
3. Back of hand................	15.53	− 8.3	10.28	14.45	13.16	
4. Forearm.....................	29.35	−55.6	20.52	18.08	20.51	
1. Forefinger..................	2.23	+25.2	4.11	3.70	4.24	Adults.
2. Cushions of hand............	6.83	+ 1.8	4.62	4.87	5.02	
3. Back of hand................	16.83	−10.3	8.73	8.40	8.82	
4. Forearm.....................	28.6	−58.8	15.40	21.75	19.30	

the forefinger, but less than the adults, and underestimate with the three other parts, whereas the adults exaggerate with the cushions of hand. Thus fewer parts of the skin lie above the indifference-point with the boys. This is one of the most important facts of the general law of evolution. Without the results for the 6-year-olds one can hardly see the full meaning of these facts, but the much greater similarity



Curve 15. 10 yrs. (G.)

Curve 16. 14 yrs. (G.)

Curve 17. Adults

Curve 18. Obs. K. 9 yrs.

Curve 19. Obs. H. 14 yrs.

between these three curves than between Curves 1–6 suggests that the change to school conditions does not prove such a contrast to pre-school conditions in the former as in the latter case. My own observations certainly justify this point, for there is no doubt that the change for the Grahamstown boys is very much greater than for the Leipzig boys. The former live a life of considerable activity and freedom in the open air whereas the latter live in a very overcrowded state in big commercial districts. It is quite possible (the different con-

ditions are doubtless very numerous) that the German boys undergo much less change in the estimation than the others, that nature's demand for a high degree of coördination in the period of early childhood is so great that it is attained in any case, and once attained, school conditions fairly closely allied to home conditions produce very little change. However, the main point of the investigation has been verified, the rise in the position of the curve with the adults.

The following table gives the averages for the *M*- and *h*-values (precision-values) for each age.

TABLE XI

	10 Years	14 Years	Adults
h-values	.05	.06	.08
M-values	13.74	11.65	9.35

Here as in Table VII.[1] the precision with the adults is greater than with either of the groups of boys. The fall in the precision from the 10- to the 14-year-olds was explained there by the oncoming of puberty. In this case it rises. The later age for puberty should explain its not falling at this point, but possibly at a later one. This great difference at the 14-year-old stage gives weight to the assumption that puberty was one of the main factors in explaining the change in the estimation.

TABLE XII

Skin-locality	Obs.	Error Per Cent.	M_o	M_u	M
1. Forefinger	K. (9 years)	+40.1	4.9	4.73	4.82
2. Back of hand	K. (9 years)	+20.7	8.84	6.81	7.9
3. Forearm	K. (9 years)	−28.1	14.4	26.34	21.22
1. Forefinger	H. (14 years)	+15.8	4.08	2.65	3.44
2. Back of hand	H. (14 years)	− 5.4	10.48	11.14	10.81
3. Forearm	H. (14 years)	−20.5	24.02	22.78	23.40

It is interesting to refer here to the experiments conducted with the two German boys, H. (14 years) and K. (9 years), brothers attending the same school, already published in the

[1] P. 277.

earlier investigation.[1] For more direct reference the points most relevant to the present discussion will be reproduced here and in curves 18 and 19.

These results are a striking corroboration of the general evolutionary law. The 9-year-old K. stands remarkably high in the estimation scale, showing exaggeration with even the back of the hand. The 14-year-old H. stands about in line with the average German 14-year-old, but exactly whether on the downward grade towards the change-point or on the upward grade beyond this to the adult stage it is not possible to say. The precision-values of .06 (K.), .056 (H.), and .08 (adults), give a very close parallel to those of the Grahamstown boys. K.'s results are undoubtedly very extreme ones for a German boy, judging from the average case.

(c) The Estimation of Large Distances, Whose End-Points Lie Near the Edges of Limbs or Other Known Boundaries

In the earlier investigation[2] on this subject, the influence of the magnitude of the distance on the correctness of the estimation was not found to be as great as Henri[3] and Fechner[4] had claimed. Clearly the whole point is such nearness of the points of the æsthesiometer to well-known boundaries as will give appreciable aid to the observer in visualizing the distance and comparing it with the more or less known size of the parts stimulated. Possibly the earlier tests did not employ large enough distances to show this influence, and further there is every likelihood that the observers strove to avoid using the auxiliary associations—they had got into the habit of avoiding all such aids in the process.

With less controlled observers this influence should rather readily show itself. It has been definitely demonstrated in the cases of the 14-year-old Grahamstown boys where a large visual distance of 45 mm. was used for the group 'back of

[1] Fitt, 'Grössenauffassung, etc.,' pp. 448 and 449.
[2] Fitt, 'Grössenauffassung, etc.,' p. 435 ff.
[3] Henri, 'Ueber die Raumwahrnehmung des Tastsinnes,' 1898, p. 61.
[4] Fechner, 'Elemente der Psychophysik,' 1860, II., p. 315 ff.

hand.' Naturally for such a large visual distance many of the touch distances must be considerably greater than 45 mm., and hence lie very near the edges of the hand. These 14-year-olds were the first investigated, and after the influence of large distances had been detected, smaller ones were used and in no succeeding case was there irregularity of this kind. The following table presents the facts.

TABLE XIII

Skin-locality	*N*	*A*	Error %	Obs.
Back of hand	28	33.75	− 20.5	Bo.
	45	50.3	− 11.8	
	15	45.5	−203.3	Br.
	45	49.13	− 9.2	
	28	28	0	P.
	45	39	+ 13.3	
	28	36.7	− 31.1	Ts.
	45	45.5	− 1.1	

One observes the same influence throughout. The additional ability to visualize, owing to the nearness of the points of a large distance to the edges of the hand, causes in each case a rise in the scale of estimation for that part of the skin. In three cases, Bo., Br. and Ts., a fairly large underestimation with a relatively small distance changes to a much smaller underestimation with a large distance, and in the case of P. correct estimation changes to overestimation. The effect of the visualizing in the first three cases is the same as the influence which the possession of a strong visual type of imagery has, viz., the estimation is more nearly correct than otherwise would be the case. This has been clearly demonstrated by the author[1] and was also noted by Washburn.[2] This generalization refers, of course, merely to a change from underestimation to more correct estimation, for Washburn's data and mine all bear on this point.

[1] Fitt, 'Grössenauffassung, etc.,' p. 444 ff.

[2] Washburn, 'Ueber den Einfluss der Gesichtsassoziationen, etc.,' in *Phil. Stud.*, XI., 1895, p. 190 ff.

Fechner's observations on this point have little or no value, owing to his unsatisfactory method and procedure.[1]

Henri finds that underestimation with small distances changes to overestimation with large ones whose end-points approach the sides of the arm. This is really the above change exaggerated and quite similar to the change in the case of P.

In the light of these facts, one must choose distances of such size as to avoid the influence of this factor, which has no relation whatever to the immediate judgment in the tactual estimation.

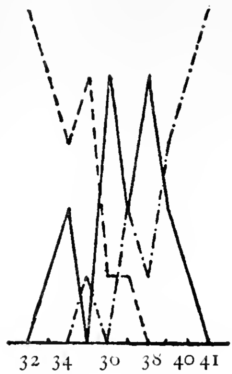
As a matter of fact the actual judgment becomes complicated by the interference of this new factor, as was brought out in the introspection of the adults.[2] This is well illustrated by the frequency-curves 20, 21, 22 and 23, of Bo. and Ts., showing the disposition of the 'greater,' 'equal' and 'smaller' judgments in the groups represented in Table XIII.

IV. Summary of Main Results

1. (*a*) The general law of tactual estimation with adults: Those parts of the skin which have very small space-thresholds overestimate two-point distances, the overestimation gradually decreasing with increase of the threshold until the indifference-point is reached where the estimation is correct. Increase of the threshold beyond this point brings underestimation, and the greater the increase the greater the underestimation.

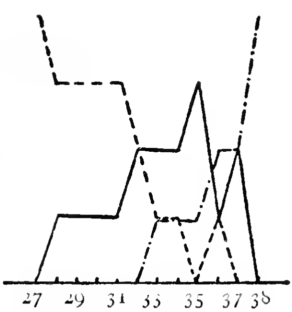
(*b*) Its biological significance: The variable law represents different degrees of adaptation of the various parts of the skin according to use. The more mobile the part of the body, *i. e.*, the greater the need of adaptation with the object of completer perception of the external world, the higher the estimation rises in the region of underestimation towards correct estimation or even higher into the region of over-estimation. The concomitant on the psychical side is the gradual increase in the degree of coördination between the pressure experience and the ability to visualize the stimulated part.

[1] Fitt, 'Grössenauffassung, etc.,' p. 452.
[2] Ibid., p. 445.

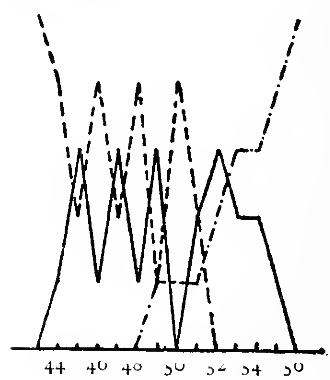


Curve 20. Obs. Ts.
Back of hand; $N = 28$;
$A = 36.7$; $M_0 = 6.85$;
$Mu = 7.26$; $M = 7.24$.

Curve 21. Obs. Ts.
Back of hand; $N = 45$;
$A = 45.5$; $M_0 = 12.17$;
$Mu = 20.23$; $M = 16.9$



Curve 22. Obs. Bo.
Back of hand; $N = 28$;
$A = 33.75$; $M_0 = 5.58$;
$Mu = 10.08$; $M = 8.39$.



Curve 23. Obs. Bo.
Back of hand; $N = 45$;
$A = 50.3$; $M_0 = 8.8$;
$Mu = 13.04$; $M = 11.35$.

"Greater" curve — · — · —
"Smaller" curve - - - - - - - -
"Equal" curve ————

2. The general law holds with boys. With the Grahamstown boys the estimation curve falls gradually from a position at 6 years, where it is only slightly lower than the average adult position, till about 13 years, where the whole curve lies considerably below the indifference-line, *i. e.*, entirely in the

underestimation-zone. Then it rises more rapidly, having crossed the indifference-line by 14 years, where the position is still slightly below the 6-year position and thus lower than the adult position.

3. The fall in the position of the curves from 6 to about 13 years of age represents a breaking-down of the touch-sight coördination due to conditions of school-activity, and the rapid rise after 13 years is attributed to the accelerating influence of puberty.

4. The precision-values of boys in this estimation are smaller than those of the adults. The 6-year-olds show on the whole remarkably high precision-values, corresponding to the frankness and immediacy of their judgments. Hence the possibility of conducting psychophysical experiments with children of this age, provided the mental processes involved are not too complex and that very little be expected on the qualitative side, in the form of direct introspection on their part.

5. Where the tactile distances are large enough to aid estimation through visualization of the stimulated part of the body, what is ordinarily underestimation changes appreciably in the direction of correct estimation, and there are indications that correct estimation or slight underestimation changes to overestimation. This additional visualization causes something of the nature of a divided judgment, appearing in a very irregular and complicated distribution of the 'greater,' 'equal' and 'smaller' decisions.

HAND–TONGUE SPACE PERCEPTION[1]

BY C. N. WATERMAN, JR.

Ohio State University

When the size of a newly formed cavity in the mouth, such as is caused by the removal of a tooth, or other dental work, is explored by the tongue, the cavity seems unusually large. This fact led to the question as to whether areal stimuli on the tongue are overestimated as compared to the same stimuli on the hand. More technically stated, we have the problem:

"What is the relation between tactual space perception on the palm of the hand, and on the tongue, when estimated in terms of visual areas?"

The following apparatus was used. In Fig. 1 is shown an ordinary chair (*A*), across the back of which was fastened the cardboard screen (*a*), extending from the seat to a distance of about eight inches below the presentation frame (*b*). In the screen a hole (*e*) was cut, through which the observer placed his right hand, resting it on the padded crossbar (*d*).

The presentation frame (*b*) was 8″ × 11″ × 14″, painted black on the inside, and fastened to the chair as indicated. At (*c*) an opening was cut through the end and bottom of the presentation frame, so that the observer sitting at (*B*) could conveniently see the visual chart (*C*) which was fastened at (*f*), and yet was prevented from getting a view of anything outside of the frame.

The chart presented at (*f*) is reproduced at (*C*), and consists of a series of twenty-seven circles, the lengths of the circumferences of which varied in steps of .25 cm. starting with a circle .50 cm. in circumference in the center, and ending with a circle 7.00 cm. in circumference at the outside. These were arranged in a spiral pattern, and lettered consecutively from 'A' to '&.'

[1] This experiment was directed by Professor A. P. Weiss, of the department of psychology.

The stimulus rack (*g*) held the set of twelve stimulus rings which were used for both hand and tongue. These consisted of a series of brass tubes, each four inches in length, and of the following circumferences in centimeters: 1.50, 2.00, 2.25, 2.50, 3.00, 3.50, 4.00, 4.25, 4.50, 5.00, 5.50, 6.00. Each was lettered to agree with the circle of corresponding circumference

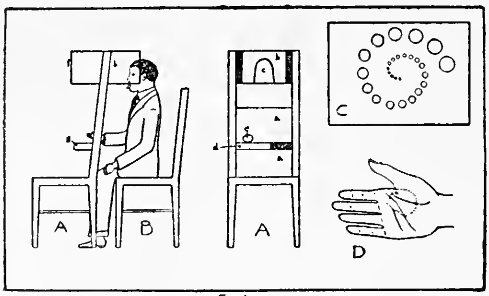
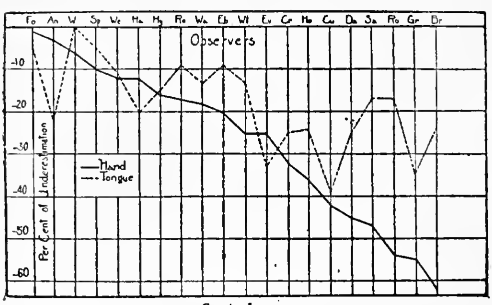
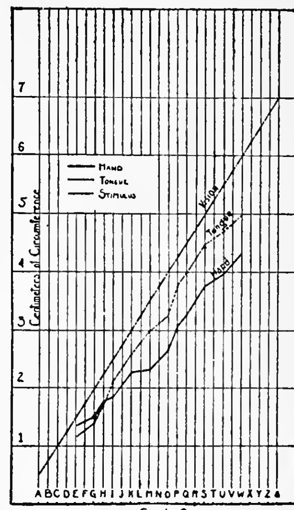


Fig 1



Graph 1.



Graph 2.

on chart (*C*), giving the series E, G, H, I, K, M, O, P, Q, S, U, W. The stimulating edges of these tubes were sharp, but smoothed so that they would not injure the tongue when applied.

When the stimulus was applied to the palm of the hand which the observer had placed through (*e*), the observer estimated and reported which of the circles before his eyes on chart *C* seemed to be the same size as the stimulus on the palm of his hand. The tongue was projected beneath the presentation frame, and the judgments and records were made in the same manner as for the hand.

On the hand, the area stimulated lay within the dotted

line shown on *D*. Upon the tongue, the area stimulated was as near the tip as possible, the stimulus being normal to the surface. This avoided strain in projecting the tongue, but care was taken that it was out sufficiently far to avoid any danger of the teeth being beneath the points stimulated, as this would have prevented uniform contact. Slight movements of the tongue were permitted, though the location of the stimulation ring upon its surface was not changed.

Each series of twelve sizes in random order was presented to the hand and to the tongue. Five of these series, no two of them alike, were presented to the observers, the order being the same for each person. Twelve stimulations of the hand were alternated with twelve stimulations of the tongue, by which procedure fatigue of the end organs was avoided to some extent at least. The full set of ten series (five series of twelve stimulations to the hand, and five to the tongue) occupied about half an hour in presentation. All the time necessary to pronounce a judgment was allowed, and the stimulus was repeated as often as requested.

In computing the data the circumference of the ring was taken as a basis. If the circumference of the stimulation ring is S and the circumference of the ring (on chart C) which the observer judges to be of equal size is V: the per cent. of over- or underestimation was derived by subtracting the circumference of the stimulation ring (S) from the circumference of the visual ring (V), and dividing the quotient by the circumference of the stimulus ring. Algebraically expressed this gives the formula $V - S/S = E$. For overestimation E is positive, and for underestimation E is negative.

Since there were twelve sizes of rings, each presented five times to the hand and five times to the tongue, the total of the E values for each observer is 120. This gives for each observer the average of sixty judgments for the hand and sixty judgments for the tongue.

The individual results are shown in Table I. Thus, for the observer Fo. the -1 in the first column means that for the hand there was an average underestimation of one per cent. on all the sizes. For the tongue the average underesti-

mation was five per cent. The average deviation for the hand was ±22, and for the tongue ±21. The results of the first two columns of this table are represented graphically in Graph 1. Twenty observers were used.

Results

Table I

Per Cent. of Underestimation

	Average		Av. Deviation	
	Hand	Tongue	Hand	Tongue
Fo	−1	−5	22	21
An	−3	−22	14	13
Wi	−6	−00	14	13
Sp	−10	−5	15	10
We	−12	−11	12	11
Ha	−12	−20	17	14
Hy	−16	−15	18	15
Re	−17	−9	14	12
Wa	−18	−13	11	16
Eb	−20	−9	10	11
Wil	−25	−13	10	7
Ev	−25	−33	10	11
Cr	−32	−25	11	11
Ho	−36	−24	13	10
Cu	−42	−39	12	18
Da	−45	−25	14	19
Sa	−47	−17	13	15
Ro	−53	−17	8	9
Gr	−55	−35	6	10
Br	−62	−24	9	14
Average	−27	−18	±13	±13

This graph shows the relationship between the estimations of hand and tongue as indicated in the first two columns of Table I. Both hand and tongue areas are underestimated when compared with the visual stimulus. For all but four observers the underestimation of the hand is greater than that of the tongue. The average underestimation of all observers for the hand is 27 per cent.; that of the tongue 18 per cent. Disregarding the absolute underestimation as compared with vision, the overestimation of the tongue as compared with the hand is nine per cent. of the circumference of

the stimulation rings. In other words, the stimulation rings presented to the tongue should be nine per cent. larger than those presented to the hand in order that the two sets may be judged equal. Rather marked individual differences are also shown on this curve.

In Graph 2 the average results of the twenty observers are presented according to the sizes of the stimuli. The dot and dash line gives the circumferences in centimeters of the visual stimuli, represented by their respective letters A, B, C, $\cdots$. Thus the circle *A* on the visual chart had a circumference of .50 cm., *B* of .75 cm. and so on. It will be remembered that of the total series of twenty-seven sizes only the following twelve were presented as tactual stimuli: E, G, H, I, K, M, O, P, Q, S, U, W, and on this graph will be found the averages of the *actual* underestimations in centimeters by the hand and tongue for each of these. For example, E has an underestimation of about .125 cm. by the hand and of about .300 cm. by the tongue.

For an explanation of the relatively greater accuracy of the tongue as compared with the hand, the following facts may be of importance. The majority of objects which are placed on the tongue and with which it comes into contact are very nearly the sizes of the stimulation rings, so that judgments made of the tongue stimulation correspond rather closely with the sizes of objects usually placed on the tongue. In other words the tongue judgments were made under the more habitual conditions, and should therefore be expected to be more accurate.

The objects which are judged by the hand, however, are as a rule larger than the stimulation rings. The palm of the hand seldom comes into contact with small objects, and even then hardly ever for the purpose of discrimination. When we wish to feel the size of anything it is moved about between the tips of the fingers. But in the palm of the hand many large and fairly homogeneous surfaces are applied as is illustrated by the grasping of the handle of a hammer, where the whole palm is stimulated equally and at the same time. Since this area is little used for discrimination and the surfaces with

which it does come into contact are large, it is not surprising that its accuracy is less than that of the tongue.

Since the problem is by no means solved by the above investigations, a more detailed and theoretical discussion will be postponed until a number of further experiments are performed. It is desired to test the influence of movement upon the judgment; the relation of the hand judgment to the tongue judgment without the visual medium, and the relationship of judgments carried out by the hand alone, and the tongue alone.

From this experiment, however, it may be stated:

1. Perception of size is more accurate with the tongue than with the hand: underestimation for the tongue 18 per cent., for the hand 27 per cent.

2. The average deviations for both hand and tongue are approximately the same.

3. The individual variations in the estimations of sizes are greater for the hand than for the tongue.

SOME AREAS OF COLOR BLINDNESS OF AN UNUSUAL TYPE IN THE PERIPHERAL RETINA

BY C. E. FERREE AND GERTRUDE RAND
Bryn Mawr College

At the psychophysical section of the first Congress for Experimental Psychology held at Giessen April 18–21, 1904, F. Schumann reported what he termed an unusual case of color blindness (his own).[1] So far as the ability to get the positive sensation is concerned, Dr. Schumann is, according to the report, totally blind to green and partially so to red.[2] In addition his case presents the following features. (1) While green light does not arouse a sensation of green, it does give red after-image and contrast sensations. Red light on the other hand gives a positive sensation but does not give either after-image or contrast sensations.[3] (2) A colorless mixed light can be matched by combining homogeneous red and green lights, but quite a little greater proportion of green is needed to give the neutral sensation than is required for

[1] Schumann, F., 'Ein ungewöhnlicher Fall von Farbenblindheit,' Bericht über den I. Kongress für experimentelle Psychologie in Giessen, 1904, pp. 10–13. See also G. E. Muller's discussion of the report, *ibid.*, pp. 20–21.

[2] His diagnosis that he is partially sensitive to red was based on two facts. (*a*) Red in the region of 670 $\mu\mu$ gave a sensation which was plainly different from yellow and could not be matched by a full spectrum gray. And (2) orange which to the normal eye appeared distinctly reddish, appeared to him a pure yellow. From this point in the orange towards the short wave-length end of the spectrum three qualities were sensed: a yellow, a blue and a band between them which could be matched with a full spectrum gray. On these facts was based the diagnosis of blindness to green.

[3] While a red light does not give green contrast sensation, it does produce an effect on a neighboring field which raises the threshold or diminishes the sensitivity to red. That is, a gray ring on a red ground appears gray but an amount of red can be added to it without being sensed which is supraliminal when red is not present in the surrounding field. In other words a physiological induction seems to be present which inhibits the complementary excitation although the induced excitation does not itself arouse sensation. We have here, therefore, another evidence that the complementary and induction relations between red and green are intact, the ability of the green excitation to arouse sensation alone being absent.

the normal eye. And (3) a yellow of the spectrum can be matched by the combination of a red and green if properly selected. In this case also a considerably greater proportion of green is needed than is required for the normal eye.

One of the writers[1] mentioned several years ago in a partial report of a somewhat extensive investigation of the color sensitivity of the peripheral retina that small areas could be found in the periphery of the normal retina showing characteristics for the colors red, green, yellow and blue similar to the Schumann case for red and green. That is, areas may be found which are totally blind or deficient to one of these colors so far as the positive response is concerned, but which seem not to be correspondingly deficient in the after-image and complementary or cancelling reactions. In fact, so far as can be told, the after-image and complementary reactions are no different in these areas from those in the immediately adjacent normal portions of the retina. We found it infeasible to test for a contrast reaction in these comparatively small and remote areas. Since that time an examination has been made of the eyes of a number of observers, more especially in connection with the work of the undergraduate laboratory, with the result that although observers may differ widely with regard to the number of the spots, their location, and the color responses affected, we are inclined to believe that the presence of such spots in the peripheral retina may be considered the rule rather than the exception.

A successful search of the peripheral retina for spots of the kind described above requires some means of making a rather minute investigation from center to periphery in a number of meridians. In our own work the rotary campimeter has been employed as a means of presenting the light to the different parts of the retina and the investigation has been made both with pigment papers and with the light of the spectrum as stimuli. In case the light of the spectrum was used, a very intensive stimulation was given. The lights were narrow bands taken

[1] Rand, G., 'The Factors that Influence the Sensitivity of the Retina to Color,' *Psychol. Rev. Monog.*, 1913, *15*, footnote, pp. 108–109. See also Ferree and Rand, 'A Note on the Determination of the Retina's Sensitivity to Colored Light in Terms of Radiometric Units,' *Amer. Jour. of Psychol.*, 1912, *23*, p. 331.

from the spectrum of a Nernst filament (corrected for impurities) operated by 0.6 ampere of current, and was of the order of intensity at the eye of 11.14×10^{-8} watt per sq. mm. for the red; 9.676×10^{-8} watt per sq. mm. for the yellow; 1.285×10^{-8} watt per sq. mm. for the green; and 0.878×10^{-8} watt per sq. mm. for the blue. The total amount of light entering the eye was in each case respectively 36.76×10^{-8} watt; 31.93×10^{-8} watt; 4.24×10^{-8} watt; and 2.90×10^{-8} watt. Lights of this intensity for the same observers were sufficiently strong, for example, to be sensed as color clear out to the limits of white light vision for all the colors but the green; and the green, it may be mentioned, could not for the observers used be made to coincide with the limits of white light vision whatever intensity of light was employed. The only effect of these greater intensities was to narrow in some cases the area of the spots previously mapped by means of the pigment paper stimuli. Thus it seems that the totally blind area is frequently bounded by a zone of weakened sensitivity. The investigation was made with three conditions of surrounding field: a gray of the brightness of the color, white and black. The only effect of the white and black surrounding fields was to increase the size of a spot mapped with a surrounding field of the gray of the brightness of the color. A preëxposure of a gray of the brightness of the color was used in each case. Care was taken to keep the intensity of the illumination of the room accurately constant throughout the investigation.

In testing for the after-image reaction, a few seconds of preëxposure was given to a gray of the brightness of the color previous to the color stimulation. The color stimulation was given for 3 sec. and the after-image projected on a gray of the brightness of the color, placed in the path of the colored light just behind the campimeter-opening. The intensity of after-image was compared as well as could be with that obtained with the same intensity and conditions of stimulation on the immediately adjacent portions of the retina. So far as could be told there was characteristically no difference between the strength of the after-image reaction in the spot itself and in

the immediately surrounding retina. In the investigation of the complementary or cancelling reaction, a combination of the complementary colors to gray was made for the surrounding portions of the retina and this stimulus was presented to the color-blind area. In each case it was seen as gray, not as the color complementary to that for which the spot was blind. We need scarcely point out that in making this test it was necessary to determine whether the amount of colored light employed would have aroused color sensation had it not been combined with its complementary color. This was done in two ways: (*a*) The amount of color used to form the combination to gray was presented to the color-blind area, combined with the proper value of a substitute sector of the gray of the brightness of the color to which the area was blind; and (*b*) the threshold was determined in the color-blind area for the color complementary to that to which the area was blind, and its value compared with the amount used in making the combination to gray. The results of both determinations showed that a true complementary action was present.

The peripheral retina spots while similar in a general way to the case described by Schumann present the following points of difference. (1) There was no detectable weakening of the sensitivity to the complementary or antagonistic color in the areas in question. And (2) no more of the color to which the area was blind was required to combine to gray with the antagonistic or complementary color than was needed on the normal areas of the retina immediately adjacent.

In the presentation of results space will be taken for only two observers, selected because of rather wide differences in the number, size and location of the spots. In Chart I. is shown results for Observer R. and in Chart II. for Observer C. The Hering pigment papers were used as stimuli and the investigation was made with surrounding field and preëxposure of the brightness of the color in each case. The limits of sensitivity were determined for each color in 16 meridians and these points are connected to give outline maps of color sensitivity for stimuli of the intensity used. In working in any of the above-mentioned meridians the stimulation was given at points

separated by no more than 1°. When a gap or a significant depression in sensitivity was found, the campimeter was rotated and the investigation made in a sufficient number of near-lying meridians to give a careful outline of the deficient area. These areas are represented on the charts in black, with letters to indicate the colors to which there is a deficiency. In case there is total blindness to the stimulus used, the spot is repre-

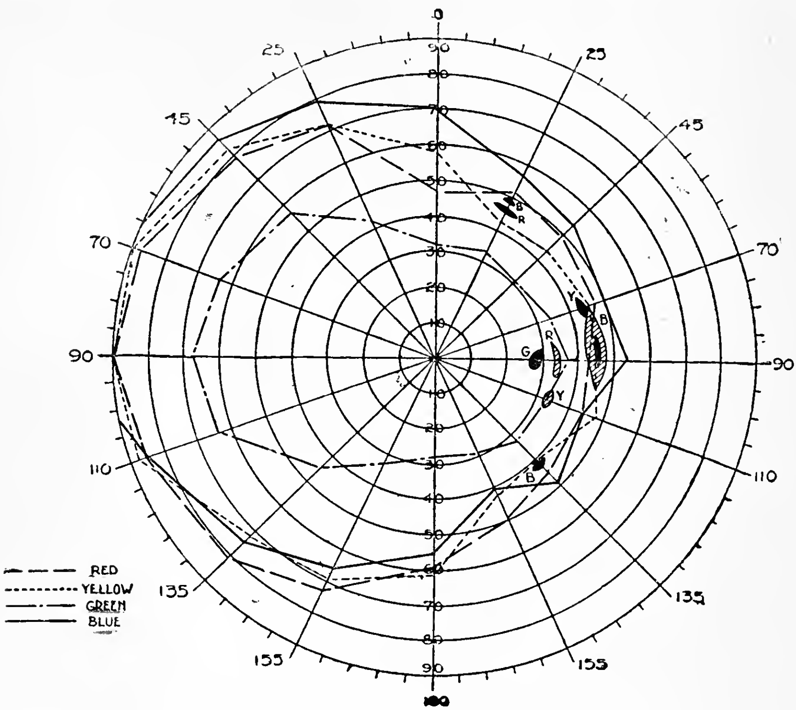


CHART I

Showing for Observer R the Areas of the Peripheral Retina having the Schumann Type of Color Blindness. These areas are represented in black, with letters to indicate the colors to which there is a deficiency. In case there is total blindness to the stimulus used, the area is represented in solid black; in case there is only a marked depression of sensitivity, the area is shaded. In this latter case areas are represented only when the depression amounts nearly to blindness. The only effect of using spectrum lights of very high intensity: red (670 $\mu\mu$), 36.76 $\times$ 10^{-8} watt at pupil of eye; yellow (581 $\mu\mu$), 31.93 $\times$ 10^{-8} watt at pupil of eye; green (522 $\mu\mu$), 4.24 $\times$ 10^{-8} watt at pupil of eye; and blue (471 $\mu\mu$), 2.90 $\times$ 10^{-8} watt at pupil of eye, was to narrow in some cases the area of the spots previously mapped by means of the pigment paper stimuli.

sented in solid black; in case there is only a marked depression of sensitivity the area is shaded. In this latter case areas are represented only when the depression of sensitivity amounts nearly to blindness. They were, for example, so insensitive that the color response could not be aroused when an unfavorable brightness of surrounding field or preëxposure was used.

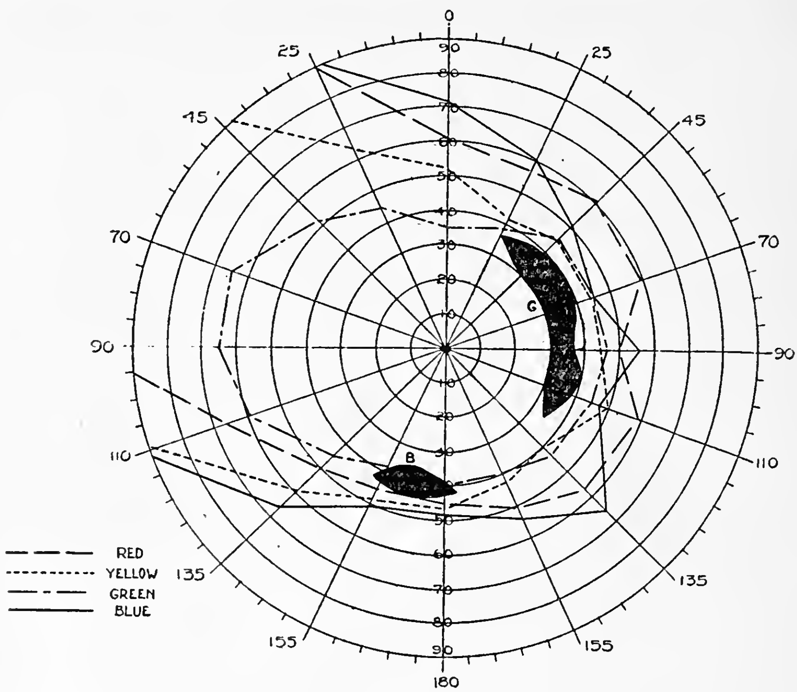


CHART II

Showing for Observer C the Areas of the Peripheral Retina having the Schumann Type of Color Blindness. The conditions of investigation and method of representation is the same in this chart as in Chart I. The results for Observers R and C are selected for presentation here because of the somewhat unusually wide difference that was found in the number, size and location of their spots.

In discussing his own case, Schumann seems to think that the phenomenon indicates that there must be more than one functional level involved in the production of visual sensation: peripheral or sub-cortical, and cortical. One of these, the peripheral or the sub-cortical, is the locus of the complementary or cancelling action, and the after-image and contrast reac-

tions. Green light in his case arouses these three reactions because the level concerned in producing them is functionally normal. Green does not arouse the positive sensation, however, because there is functional deficiency in the remaining level or levels. G. E. Müller, who also made supplementary tests and experiments on Schumann and discussed Schumann's report at the Congress in Giessen, concurs strongly in the conception that more than one functional level is needed to explain the Schumann case. At this same Congress Müller[1] discusses seven types of color-blindness which he further believes can be explained best on the conception of more than one specific functional level, the processes of which may be separately deficient. His theory of color vision is in fact here elaborated to include both peripheral and central visual processes. While we have no wish to engage in theoretical discussions at this stage in our own work, it may not be out of place, in addition to the results presented in this paper, to call to mind in this general connection our experiments of the effect of the achromatic excitation on the chromatic which seemed to indicate very strongly that this effect both quantitative and qualitative takes place at some level posterior to that of the cancelling, after-image, and contrast reactions.[2] We have also obtained other results, as yet unpublished, on contrast induction in the far periphery of the retina which seem to indicate that the deficiency which for these portions of the retina prevents the color stimulation from producing sensation is in part at least posterior to the level at which induction takes place. Moreover, unless the complementary action were intact in case of

[1] Müller, G. E., 'Die Theorie der Gegenfarben und die Farbenblindheit,' Bericht über den 1. Kongress für experimentelle Psychologie in Giessen, 1904, pp. 6–10. The conception of a central deficiency was used as early as 1868 by Niemetschek to explain color blindness (*Prager Vierteljahrschrift*, *100*, p. 224).

[2] The results of these experiments have as yet been published only in part. (See "An Experimental Study of the Fusion of Colored and Colorless Light Sensations. The Locus of the Action," *Journ. of Philos., Psychol. and Scientific Methods*, 1911, *8*, pp. 294–297.)

The fuller publication has been delayed for one reason because we have felt the need of giving a somewhat exhaustive résumé of the work that has yet been done on the subject. This work has been so scattered and unsystematic and so much of it appears hidden under titles that give no indication that it is there, that the work of compiling has been somewhat time consuming.

blindness to one color over a whole or part of the retina, it would seem that white light should always be sensed by the subject in the tone complementary to that for which the blindness exists.[1] Or to put the matter more conservatively in terms of color-blindness testing, it would seem impossible ever in such cases to match a full spectrum gray to the color to which the subject is blind, or to all combinations of complementary colors, or to any in fact except the pair to one member of which the defect exists.

In the study of the eye as a recording instrument it is always a helpful feature to take our start from the physical recording instruments which respond to light. Not only are the characteristics of a given one of these instruments more accessible to study than the eye, but the instrument itself can be changed and the effect produced be noted. Also different types of instrument are accessible to study. Just as the simpler work on the study of the physical instruments serves as a helpful methodological guide in the experimental determination of the characteristics of response of the eye, so may we get methodological helps from the study of these instruments which may be of service in forming our conceptions of the actions and functional relations of the cerebro-retinal structure.

One of the characteristics of the instruments which respond to radiant energy is with the exception of the photographic plate a surface or layer in which the energy of the light-wave is transformed into an effect which it is the function of another part of the apparatus to record. This transformation, moreover, in case of some of these instruments: the selenium cell, the photoelectric cell and the photographic plate, is selective;

[1] It is possible that an alternative explanation of the above point may be found in some of the other types of modification of existing theories, *e. g.*, Fick's and Leber's modification of the Helmholtz theory to explain the variations in the color sensitivity of the peripheral retina (see Fick, A., 'Arbeiten aus dem physiol. Laborat. der Würzburger Hochschule,' pp. 213–217; Leber, T., *Klin. Monatsblätter f. Augenheilk.*, 1873, *11*, pp. 467–473). The conception of different functional levels, however, should at least be recognized as one of the possibilities of explanation for monochromatic deficiencies, and as a very plausible and perhaps necessary assumption to explain phenomena of the kind described by Schumann and by the present writers in this paper.

i. e., it is different in amount for the different wave-lengths[1] and the selectiveness varies with the intensity of light used. There is an analogue to this in the selectiveness of the achromatic response of the eye to wave-length and the variation of this selectiveness with change of intensity of light. (In case of the achromatic response this is known as the Purkinje phenomenon.) Some of these instruments also show like the eye a lag in coming to their maximum of response, a fatigue effect, and an after-effect. All of these effects are characteristic of the receiving part of the instrument, not of the recording mechanism. The final form into which the response of the instrument is put is, however, a function of the recording part of the apparatus. Moreover, either one of these parts of the apparatus may, we scarcely need point out, be separately deficient without the impairment of the other. In view of the general similarity in characteristics of response shown between the eye and these instruments, it would not seem entirely unreasonable to suppose,[2] therefore, even in advance of a decisive amount of the evidence which seems to be accumulating, that the visual apparatus consists of receiving and recording parts or levels both of which are necessary to the final response, but either of which may be separately deficient without impairment of the function of the other.

[1] Griffith, I. (*Phil. Mag.*, 1907, *14*, (6), p. 297) working with ultra-violet radiations and Dember, H. (*Ber. d. kgl. sächs. Akad. d. Wiss.*, 1912, *64*, p. 266) working with the visible spectrum both claim that the photoelectric cell is selective in its response to intensity of light. Elster and Geitel (*Phys. Z.*, 1913, *14*, p. 741; 1914, *15*, p. 610) however, found a constant relation between intensity of light and the response of the cell except in case of very intense light. For a fuller discussion of this point see Ferree and Rand, *Psychological Review Monographs*, 1917, XXXII, (5), p. 46.

[2] It will be understood that an analogy between the eye and the physical recording instruments is attempted here primarily as illustrative rather than as argumentative.

THE PROGRESSIVE ERROR OF THE SMEDLEY DYNAMOMETER

BY SAMUEL C. KOHS

Buckel Foundation, Stanford University

Introduction

There are, undoubtedly, scores of Smedley dynamometers now in actual use in this country. That these instruments may be furnishing inaccurate data, inaccurate to all who may wish to make use of such measurements, has already been intimated by a few. At the present time it would hardly be safe procedure to accept any table of figures at its face value. The reason for this we hope to demonstrate in the present article. When the same amount of pressure yields as much as ten or fifteen pounds difference in the registrations of two different dynamometers, the situation is scientifically serious enough for thoughtful consideration. It is to be regretted that but little space has yet been devoted in the literature of tests to an adequate discussion of errors of instruments. In the meantime we have been accumulating numerous records of dynamometer reactions which, from a rigorously scientific viewpoint, are questionable. Strictly speaking, the machine, following natural laws, is honest. It is the injected human element which makes for error. In the interests, therefore, of accuracy and uniformity this study was undertaken.

Previous Literature: Psychological

Three contributions may here be mentioned.

Wallin in 1912[1] attempted to determine the variability of his instrument and found the following: (1) that the spring was too "stiff"; (2) that as the pressure increased so did the margin of inaccuracy; and (3) that different days (in December and in March) yielded different registrations for the same

[1] 'Experimental Studies of Mental Defectives,' Baltimore: Warwick and York, p. 155.

strain. For the first two conclusions a table of data is published, for the third, however, no figures are presented.

Whipple in 1914[1] advises that the calibrations be checked occasionally to eliminate the possibility of instrumental error.

Mead in 1916[2] used a two-spring Narragansett Machine Company dynamometer. Upon testing his instrument he found that it registered too low. Utilizing one standard weight, 32.5 pounds, his registrations read: 28, 28, 28, 28 when the weight was placed on top. And when hung below, in the 13.5 ounce carriage, the readings were 24, 23, 25, 25.

Previous Literature: Physical

The treatment of the subject in textbooks on mechanics is meager and fragmentary. Here and there one finds mention of the dynamometer as an example of the possible practical application of the laws discussed under the general heading of elasticity. Those who wish further information regarding this phase of the problem may find additional items of interest in the references cited at the end of this article under "Physical Aspects."

The principle underlying the working of the dynamometer is this: Below the elastic limit of a given material (both for compression and extension) the stretch (strain) is proportional to the load (stress) producing it. Presented in another form: elasticity is equal to the stress divided by the strain. And since elasticity is constant for our instrument, the amount of compression will always be proportional to the amount of force applied. The bearing of this important fact upon the recording of dynamometer readings will be made evident in a moment.

Present Series of Tests

On December 28, 1916, between two and three in the afternoon, the Buckel Foundation dynamometer was submitted to

[1] 'Manual of Mental and Physical Tests,' Vol. I., Baltimore: Warwick and York, p. 365.

[2] 'Relations of Intelligence to Mental and Physical Traits,' Columbia University Contributions to Education, No. 76, p. 117.

the following ordeal: After tightening the few screws (front and back) we tied the instrument securely to a horizontal support, with the scale visible and with the outer stirrup nearest the floor. We then threw a rather stout piece of cord (doubled) over the center of the inner stirrup so that the cord hung on both sides of the outer stirrup. There was contact with the cord at these two points, the friction of which we eliminated by thorough greasing. The dangling ends were then knotted, our weights being suspended from that place. In all the trials the weights were very carefully hooked on the cord, thus keeping our measuring force constant. For each trial the weights were removed, the registering dial adjusted at 0 kg., the proper number of weights suspended and the reading taken from the upper or right-hand edge of the dial. The stationary dial was moved out of the way, past the 100 kg. mark.

In the following table are represented the registrations for the different standard weights, beginning with five kilograms. (Visualized in Graph I.)

TABLE I

DYNAMOMETER READINGS PER STANDARD WEIGHT (STANFORD)

Standard Weight (Kg.)	Dynamometer Readings						Average Error	Increment of Error
	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Average Reading		
5	4.25	4.5	4.5	4.5	4.5	4.45	−0.55	+0.55
10	9.0	9.25	9.5	9.25	9.5	9.30	−0.70	+0.15
15	13.5	14.0	14.0	13.75	14.0	13.85	−1.15	+0.45
20	18.0	18.0	18.0	17.5	18.0	17.90	−2.10	+0.95
25	22.8	22.75	23.0	23.0	22.75	22.86	−2.14	+0.04
30	27.0	27.0	27.0	27.5	28.0	27.30	−2.70	+0.56
35	32.0	32.25	32.2	32.0	32.75	32.24	−2.76	+0.06
40	36.25	36.5	36.75	36.75	36.9	36.63	−3.37	+0.61
Average increment of error								+0.42

We note that the mean increment of error is equal to +0.42 kg. for every 5 kg. of standard weight. And, therefore, the differences between standard weight and dynamometer reading increase, for every 5 kg., in arithmetic progression.

Having ascertained this fact, we may now correct Table I. for errors of observation, yielding Table II.:

TABLE II

TABLE I CORRECTED FOR ERRORS OF OBSERVATION

Stsndard Weight	Dynamometer Reading	Error	Increment of Error
5	4.58	−0.42	+0.42
10	9.16	−0.84	+0.42
15	13.74	−1.26	+0.42
20	18.32	−1.68	+0.42
25	22.90	−2.10	+0.42
30	27.48	−2.52	+0.42
35	32.06	−2.94	+0.42
40	36.64	−3.36	+0.42
45	41.22	−3.78	+0.42
50	45.80	−4.20	+0.42
55	50.38	−4.62	+0.42
60	54.96	−5.04	+0.42

Error index = 0.916,
Correction constant = 1.09.

The index of error is a constant and is equivalent to the dynamometer reading per standard weight divided by standard weight. Thus in Table II., for the standard weight 5, the dynamometer reading is 4.58. The error index is, therefore, equal to 4.58/5, or 0.916. This remains constant no matter how small or how large the force exerted upon the instrument.

TABLE III

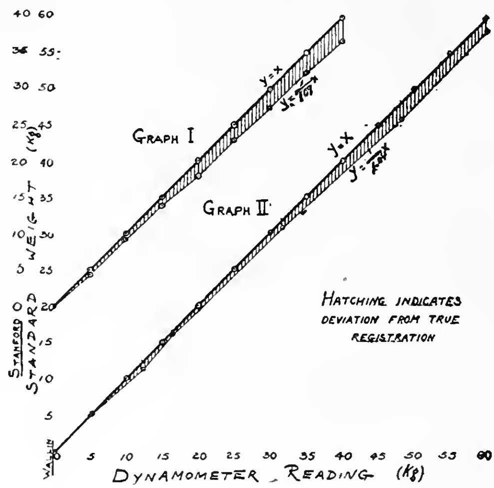
CORRECTIONS FOR DYNAMOMETER READINGS OF TABLE II

Dynamometer Reading (Kg.)	Positive Correction (Kg.)	Dynamometer Reading (Kg.)	Positive Correction (Kg.)
5	+0.45	45	+4.05
10	+0.90	50	+4.50
15	+1.35	55	+4.95
20	+1.80	60	+5.40
25	+2.25	65	+5.85
30	+2.70	70	+6.30
35	+3.15	75	+6.75
40	+3.60	80	+7.20

We may also obtain a correction constant by dividing the standard weight by its dynamometer reading. Thus in Table II. for the standard weight 10 the dynamometer reading is 9.16. To obtain the correction constant, 10 is divided by 9.16, yielding a quotient of 1.09. This factor remains constant no matter what the effort exerted. Thus a reading of 16 kg. really signifies the exertion, on the part of the subject, of 17.44 kg. of effort (16 kg. × 1.09 = 17.44 kg.). And a reading of 57.5 kg. really signifies the exertion of 62.7 kg. of

effort. A difference between the two, in the first case of 1.44 kg. (3 lbs. $2\frac{3}{4}$ oz.), and in the second of 5.2 kg. (11 lbs. $7\frac{2}{5}$ oz.). The differences certainly are serious enough to emphasize the necessity for correction. In Table III. are noted the corrections for the different dynamometer readings.

It is evident that the numerical value of the correction for each reading is the same as the numerical value of the error, except that one has a positive and the other a negative algebraic sign. To obtain the true force exerted, the correction quantity is added to the dynamometer reading. The amount of error indicates how far the dynamometer falls short of the true registration.



The algebraic expression of the line indicating the condition in which standard weights and dynamometer readings are in agreement is $y = x$ (perfect correlation). Should there be a progressive disagreement, as in our case, the correction constant being 1.09, the above-mentioned algebraic expression becomes $y = \frac{1}{1.09}x$ ("x" standing for dynamometer reading, and "y" for standard weight).

Wallin's Series of Tests

Treating Wallin's data similarly, Table IV. is obtained, in which the registrations per standard weight are recorded. Presented in curve form Graph II. is obtained.

Table IV

Dynamometer Readings per Standard Weight (Wallin)

Standard Weight	Dynamometer Reading[1]	Error	Increment of Error
5.2	5.2	0	——
12.4	11.7	−0.7	+0.7
16.5	16.4	−0.1	−0.6
31.7	30.6	−1.1	+1.0
48.2	45.7	−2.5	+1.4
60.3	58.2	−2.1	−0.4

[1] Averages based on from 3 to 20 trials per standard weight.

From the readings of Graph II., abscissas and ordinates, Table IV. may be converted into Table V.:

Table V

Dynamometer Readings per Standard Weight (Wallin)

Standard Weight	Dynamometer Reading	Error	Increment of Error
5	5.0	0	——
10	9.5	−0.5	+0.5
15	14.6	−0.4	−0.1
20	19.5	−0.5	+0.1
25	24.1	−0.9	+0.4
30	28.8	−1.2	+0.3
35	33.6	−1.4	+0.2
40	38.1	−1.9	+0.5
45	42.7	−2.3	+0.4
50	47.5	−2.5	+0.2
55	52.6	−2.4	−0.1
60	57.7	−2.3	−0.1
Average increment of error.			+0.19

Since the average increment of error is +0.19, we may again state Table V., now corrected for errors of observation. This is done in Table VI.:

Table VI

Table V Corrected for Errors of Observation

Standard Weight	Dynamometer Reading	Error	Increment of Error
5	4.81	−0.19	+0.19
10	9.62	−0.38	+0.19
15	14.43	−0.57	+0.19
20	19.24	−0.76	+0.19
25	24.05	−0.95	+0.19
30	28.86	−1.14	+0.19
35	33.67	−1.33	+0.19
40	38.48	−1.52	+0.19
45	43.29	−1.71	+0.19
50	48.10	−1.90	+0.19
55	52.91	−2.09	+0.19
60	57.72	−2.28	+0.19

Error index = 0.962,
Correction constant = 1.04.

In Table VII. are noted the corrections for the readings of this dynamometer:

Table VII

Corrections for Dynamometer Readings of Table VI

Dynamometer Reading (Kg.)	Positive Correction (Kg.)	Dynamometer Reading (Kg.)	Positive Correction (Kg.)
5	+0.20	45	+1.80
10	+0.40	50	+2.00
15	+0.60	55	+2.20
20	+0.80	60	+2.40
25	+1.00	65	+2.60
30	+1.20	70	+2.80
35	+1.40	75	+3.00
40	+1.60	80	+3.20

A Comparison of the Accuracy of the Two Dynamometers

In comparing Tables III. and VII., it will be apparent that the error of the Stanford dynamometer is slightly more than twice the error of Wallin's instrument. Thus a reagent exerting exactly 50 kg. of force upon both machines would register 45.80 kg. on one (Stanford), and 48.10 kg. on the other (Wallin). Or, stated differently, the same registration of 50 on both machines would signify the exertion of a force equal to 54.50 kg. on the Stanford instrument, and only 52.00 kg. on the Wallin instrument. (A difference of 2.5 kg. or 5½ lbs.) The greater the amount of force exerted, the greater will be the disparity between the registrations of both dynamometers.

The folly, therefore, of comparing tables of uncorrected dynamometer readings is, no doubt, evident. The same holds true, of course, when individuals, low and high, of a homogeneous group are compared.

Tabulating the accuracy differences of both dynamometers we obtain Table VIII.:

Table VIII

Accuracy Differences of the Two Dynamometers

	Error Index	Correction Constant	Curve	Force 10 Kg.	Force 30 Kg.	Force 60 Kg.
Stanford instrument.........	0.916	1.09	$y = \frac{1}{1.09}x$	9.16	27.48	54.96
Wallin instrument...........	0.962	1.04	$y = \frac{1}{1.04}x$	9.62	28.86	57.72

Conclusions

1. Unless corrected, the dynamometer readings do not accurately register the force applied.

2. The error ratio is a constant and may easily be determined.

3. The amount of error increases by arithmetic progression as the force exerted increases.

4. Error ratios (error indices) of dynamometers vary. Therefore the error index of each, and the correction constant for each should be determined.

5. The manufacturers of dynamometers should either furnish accurately calibrated instruments, or mark each machine with its index of error and its correction constant.

6. For purposes of accuracy, dynamometers should be checked for correction at least once every three months. Quoting Auerbach: "Natürlich müssen die Dynomometer vor dem Gebrauch durch bekannte Kräfte geaicht werden, und diese Eichung ist öfters zu wiederholen, da die elastischen Kräfte bekanntlich nicht unbeträchtlichen zeitlichen Änderungen unterliegen" (p. 140). ("Of course dynamometers should be checked with known forces before being used, and this check should be oft repeated, for, as is known, elasticity is significantly affected by temporal changes.")

7. Tabulations of dynamometer reactions should be corrected before publication.

8. It is recommended that some standard procedure for correction be followed by all.

Bibliography

Physical Aspects

1. Auerbach, F. Messung von Massen und Kraften. (In A. Winkelmann: Handbuch der Physik, Erster Band, Erste Halfte, pp. 123–140. Leipzig: Barth, 1906, 544 pp.)
2. Daniell, A. *A Text Book of the Principles of Physics.* Pp. 36–38. London: Macmillan, 1885, 688 pp.
3. Müller-Pouillet. *Lehrbuch der Physik und Meteorologie.* (Revised by L. Pfaundler.) Erster Band, pp. 334–341. Braunschweig: Viewig u. Sohn, 1886, 888 pp.
4. Whiting, H. *A Course of Experiments in Physical Measurements.* Pp. 337–340. Boston: Heath, 1891, 1222 pp

Psychological Aspects

5. MEAD, C. D. Relations of Intelligence to Mental and Physical Traits. Columbia University Contributions to Education, No. 76, 1916, 117 pp.
6. WALLIN, J. E. W. *Experimental Studies of Mental Defectives.* Baltimore: Warwick and York, 1912, 155 pp.
7. WHIPPLE, G. M. *Manual of Mental and Physical Tests. Simpler Processes.* Baltimore: Warwick and York, 1914, 365 pp.

DISCUSSION

THE TILTING BOARD AND ROTATION TABLE

In the JOURNAL OF EXPERIMENTAL PSYCHOLOGY for October, 1916, Professor Warren and Dr. Reeves describe a combined tilting board and rotation table. It may be of some interest to know that the writer has used successfully an even simpler device for supporting the tilting board and rotation table combined.

The board as described by the authors is to be provided with six ring bolts, one in each corner and one in the middle of each of the long sides. Four pieces of rope, about six feet long (the length depending upon the height of the ceiling), have spliced into one end of each a snap hook, obtainable at any harness makers (preferable to an open hook, as it avoids accidents); the other ends of the rope are spliced into a large ring, this ring is suspended from a stout hook placed in the ceiling, by means of a rope of sufficient length to bring the board at the desired height. When used as a tilting board, two of the ropes are hooked into the rings at the side, the other two being tied up out of the way; used as a rotating table, the four ropes are hooked into the four corners of the board. It is convenient to have the rope that suspends from the ceiling hook supplied with an extra ring, somewhere near the middle, so placed that when the apparatus is hung up, the board, when in a vertical position, just clears the floor. If new rope is used it should clear perhaps by three inches to provide for stretching later. The full length of this ceiling rope may then be so adjusted that when the board is used as a rotation table it may be within a foot or two of the floor which renders it easy for the subject to reach and removes the fear, that some subjects sometimes have, of trying the experiment if the board is up high.

Suspended by the two ropes at the centers of the sides of the board, the apparatus is as delicate as though it had knife-

edge bearings and it may thus be used for the experiment proving that mental activity induces the flow of blood to the brain. For this experiment, supports a little shorter than the height of the board are placed under each end (small tables answer well) so that the board can have a movement from the horizontal of only a few inches. Spring balances of suitable sensitivity are fastened to each end of the board and also to the table; these will control the oscillation of the board due to the breathing. The tilting of the board may be magnified and made visible to a class, by the use of a light indicator in the form of a lever with the fulcrum rather close to the end near the board and adjusted to touch the board on the under side, so that as the end goes down, even very slightly, the other end goes up rapidly.

HENRY H. GODDARD

Journal of Experimental Psychology

Vol. II, No. 5. October, 1917

A REPETITION OF EBERT AND MEUMANN'S PRACTICE EXPERIMENT ON MEMORY[1]

BY H. B. REED

University of Idaho

Reviewers of Ebert and Meumann's practice experiment upon memory are generally agreed that the experiment must be repeated before their conclusion can be accepted as valid for the facts of the experiment. Lack of space prevents me from taking notice of the reviews and criticism that have been made on their work. The references of Müller, Sleight, and Thorndike cited below will be found valuable in this connection.

Suffice it to say that the general conclusion drawn from the results was that a special training of the memory improves it in general. But since the amount of transfer from the special training to the tests was not equal in all the tests, a particular conclusion was drawn in that the amount of transfer follows the law "that the special memories are improved exactly to the degree that the learning materials and learning methods are related to the specially trained memory" (p. 200). These conclusions have not been accepted and have been criticized among others by G. E. Müller[2] and R. Wessely[3] in Germany, by W. G. Sleight[4] in England, and W. F. Dearborn[5] and E. L. Thorndike[6] in America.

So far as is known to the writer, Meumann took no notice of his critics except G. E. Müller whom he ridicules for his

[1] *Archiv. f. ges. Psychol.*, 4, 1904, 1–232.

[2] *Zeit. f. Psychol.*, 39, 1905, 111–125.

[3] *Neues Jahrb. f. Päd.*, 16, 1905, 296–309; 371–386.

[4] *British Jour. Psych.*, 4, 1911, 386–456.

[5] *Psych. Bull.*, 6, 1909, 44.

[6] 'Psych. of Learn.,' 1913, 369–376.

overemphasis upon "precision." He states that if his own explanation of transfer is mysterious, that of Müller[1] is equally so. In confirmation of his conclusions, he refers to the experiments of Winch,[2] Daniel Starch,[3] Coover and Angell,[4] and Wallace Wallin,[5] which he reviews as the 'Pädagogisch wichtigste.' Nowhere does he refer to the experiment of Thorndike and Woodworth[6] which was published three years before his own, although the remarks upon special memories in the introduction of the report of his experiment have a striking similarity to the opening remarks of Thorndike and Woodworth upon the same subject. Nor does he take notice of other experiments reaching different conclusions, *e. g.*, those of James,[7] Fracker,[8] Klein,[9] Foster,[10] Ruger,[11] and Sleight,[12] all of which are reviewed in a book to which he makes more than one favorable reference.[13] It appears, therefore, that Meumann never retracted the fundamental conclusions of his experiment.

Since this experiment of Meumann forms the background of most that is said in his 'Psychology of Learning,' and of much that is said in his 'Vorlesungen,' and has elsewhere occupied a prominent place in psychological literature, and coming as it does from the hand of one of the foremost psychologists of his time, it appeared to the writer that it ought to be repeated, and that in this way many of the important criticisms that have been made upon it might be answered. I shall refer to this repetition as the test experiment which I now wish to describe.

There were two series of tests: I., memory-span tests, and, II., learning tests.

[1] *Vorlesungen*, III., 1914, p. 143 ff.

[2] *British J. of Psych.*, 2, 1908, 284–293; 3, 1910, 386–405.

[3] *J. of Ed. Psych.*, 3, 1912, 209–213.

[4] *Amer. J. of Psych.*, 18, 1907, 237–340.

[5] *J. of Ed. Psych. Mon.*, 1911.

[6] *Psych. Rev.*, 8, 1901, 247–261; 384–395.

[7] *Principles of Psych.*, 1, 1893, 667.

[8] *Psych. Rev. Mon. Suppl.*, 4, 1908, 56–102.

[9] Klein, Thorndike, 'Psych. of Learn.,' 401 ff.

[10] *Jour. of Ed. Psych.*, 2, 1911, 11–21.

[11] *Archives of Psych.*, No. 15, 1910.

[12] *British Jour. of Psych.*, 4, 1911, 386–456.

[13] Thorndike's 'Psych. of Learn.'

I. The memory-span tests consisted of:

1. Thirteen double series of consonants ranging from five to seventeen in the first test. In the second, the quantity was increased to twenty letters.

2. Sixteen double series of numbers ranging from five to twenty. In the second test, the largest quantity was seventeen.

3. Eight double series of nonsense syllables ranging from four to eleven.

4. Twelve series of disconnected one-syllable words ranging from four to fifteen.

5. Nine series of Latin-English vocabularies ranging from three to eleven pairs, tested by the Treffer method.

6. Passages from Wordsworth's 'She was a phantom of Delight,' The 'Solitary Reaper,' 'Stepping Westward,' and 'To a Highland Girl,' with the following number of words per passage in the first test: 12, 14, 16, 18, 21, 22, 24, 26 and 28. In the second test, the quantities were: 11, 14, 16, 18, 20, 22, 24, 26, and 28.

7. Eleven passages of prose from Locke's 'Essay,' ranging from twelve to thirty-four words.

The first four of the above tests were presented in an auditory form, at the rate of one member a second. The last three tests were also read aloud at a slow rate, but not at the rate of one member per second. The reproduction in all cases was written upon paper by the subject immediately after the stimulus ceased.

II. The learning tests were all presented visually and were relearned after twenty-four hours. They consisted of:

8. Three series of nonsense syllables, consisting of ten, twelve, and sixteen syllables respectively. The syllables were made from the English alphabet, and according to the conditions of Müller and Schumann.

9. Two series of meaningless visual diagrams of twelve members each, the same as those used by Ebert and Meumann.

10. Thirty Latin-English vocabularies.

11. Forty Latin-English vocabularies.

12. Two eight-line stanzas of poetry selected from Wordsworth's 'To the Daisy.'

13. A ten-line passage from Locke's 'Essay.'

The experimenter's function in the learning test was to see that the subject carried out his instructions, and to keep account of the number of repetitions made and of the time, which was taken with a stop watch. The time was measured from the beginning of the test to the end of the first correct reproduction. The same measurements were also kept for the syllable rows in the practice series. In the calculation of the data, only the times were taken into account.

The test experiment was made with fourteen adult subjects, but later reduced to thirteen, at the University of Wyoming during the summer school session of 1915. Eight of these constituted Group I., and the other six, later reduced to five, constituted Group II.

The eight subjects of Group I. did both the practice and the test series of the experiment while the six of Group II., later reduced to five, omitted the practice series but did the test series at the same time as Group I. Group II. was the control group. The two groups were divided with respect to equal learning ability as accurately as this could be determined by the experimenter's personal judgment, but, as a matter of fact, the average learning ability of Group II. was slightly superior, which makes the results for control purposes all the better.

Like the experiment of Ebert and Meumann, the test experiment had a practice series which consisted of learning and relearning twelve-syllable nonsense series, constructed as explained above. Those which were learned one day were relearned twenty-four hours later. The form of presentation was somewhat different from that in Test 8. Instead of the experimenter allowing the subject to reproduce the syllables at his own rate, he kept presenting them in a mechanical way at the rate of one per second until the learner succeeded in reproducing for the first time each syllable before it was turned. The series was then considered learned. The four methods of learning: the whole, the part, and the two mixed methods, followed by Ebert and Meumann's subjects, were also followed by the subjects of Group I. of the test experiment. The order

of the different learning methods was varied irregularly in the test experiment. Ebert and Meumann kept the quantity of learning material constant for each practice exercise so that each subject learned two series of syllables daily. But in the test experiment, the constancy was put upon the amount of time. This was limited to thirty minutes for each exercise, except when the subject failed to learn one twelve-syllable row. In that interval, each subject relearned what he had learned the day before, and then learned as much new material as possible in the space of time remaining. Four of the subjects in Group I. had two practice exercises daily, one in the morning and one in the afternoon, while the other four had only one practice exercise daily in the morning. Several of the subjects omitted Sundays and other holidays. Approximately fifteen days were spent in practice by Group I., while upon the five days immediately before and after the practice, Groups I. and II. took the tests described above. The test experiment may therefore be considered as repeating only the first practice period of Ebert and Meumann together with the tests that came before and after that period. If this part of the experiment is confirmed or refuted, the writer is of the opinion that a repetition of the remaining parts would not essentially change the character of the results.

The course of improvement in the training series is represented in the practice curves of Fig. 1 and Fig. 2. The points in the ordinates are based upon the average learning and relearning times per series for each day of practice. That the character of the improvement in learning nonsense syllables, which is memorial or ideational in function, is of the same sort as that of learning telegraphy or typewriting and of other sensori-motor functions is evident from the character of the curves, which are typical learning curves. The individual differences, however, are striking.

The relearning curves either vary about a horizontal line or else run closely parallel to the learning curves. Contrary to our expectation from the results of Ebbinghaus, Müller, and Schumann, Radosawljevitch on the relation of the amount saved to the strength of the imprinting, a high learning time

is in most cases not followed by a correspondingly low learning time, but rather the converse. But it should be said that no special light was thrown upon the causes of improvement. In the cases under observation, they appeared not to consist in lack of effort or interest, or the development of bad learning methods such as unusual pronunciation. Nor did the coöperation of the will appear to be of absolute and controlling importance, as it is stated by Meumann, for no exertion of the will prevented bad records or plateaus. See his 'Psychology of Learning,' pp. 64, 66, 74, 115, 139, 283, 287, and 361.

But whatever may be the cause of improvement, the curves of Fig. 1 and Fig. 2 show that a very great improvement was made by the subjects of the test experiment in learning nonsense syllables. The question now is: How much of this improvement was transferred to other mental functions?

The measures employed by Ebert and Meumann to determine the effect of the improvement gained in learning nonsense syllables upon the memory-span functions were the memory-span, the first excess of one third error, and the first excess of one half error. The measure of the memory-span was taken as the number of members in the largest series correctly reproduced in any given test. The amount of error was calculated as follows: "The omission or addition of a member was considered as 4/4 error, a displacement in the series of more than one position as 3/4 error, and the displacement of the same sort of exactly one position as 2/4 error and a correction as 1/4 error (*op. cit.*, p. 11)."

In the test experiment, the three above measures were used in the same way. But to increase the reliability of the tests five additional measures were used as follows:

1. The *method of average error* per quantity on series in each test calculated according to the scale of Ebert and Meumann for each series. From the results thus obtained, the average was calculated for all the series in a test together with the average deviation from the individual series.

2. The *method of retained members*, according to which the number of members correctly reproduced regardless of order and position was calculated in each series of a test. From

these results the average and the average deviation were calculated as in 1.

3. The *method of retained groups*, according to which the largest number of members reproduced that had the same

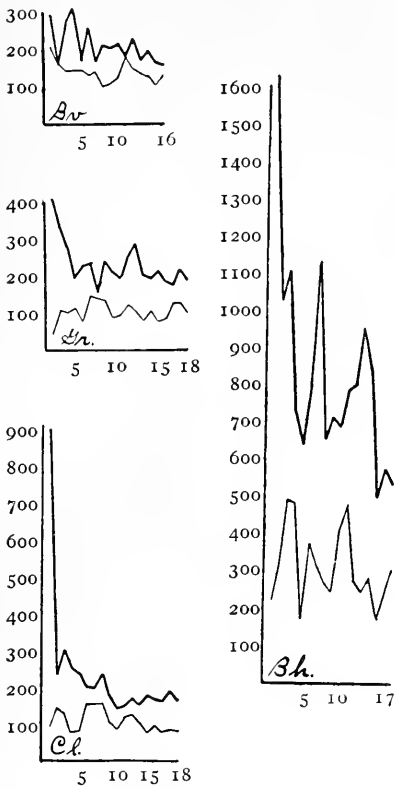


FIG. 1.

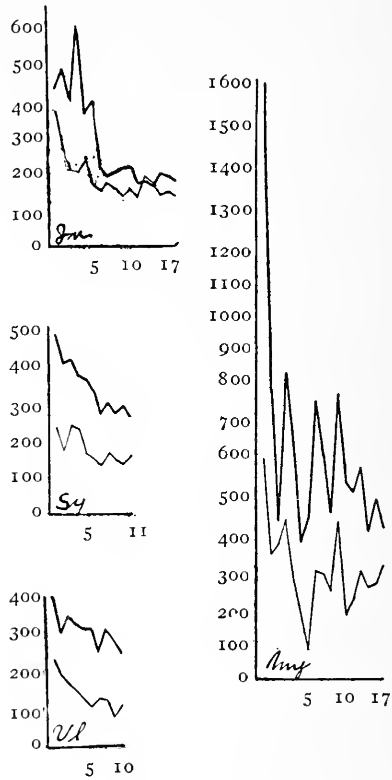


FIG. 2.

FIGS. 1 AND 2. Practice Curves Showing Course of Improvement in Learning and Relearning Twelve-syllable Rows of Nonsense Syllables. The time in seconds is plotted along the ordinates and successive practice days along the abscissas. The heavy line above represents the average learning time per series from day to day; and the light line below, the corresponding relearning time.

order of succession as those in the original series was calculated for each series in a test. The average and average deviation were determined as in 1.

4. The *method of retained positions*, according to which the number of members in the reproduction that had the

same position relative to either the first or the last end of the series was calculated. The average without the A.D. was determined as in 1.

5. The method of correlation by the Pearson Coefficient, the application of which will be described later.

In Table I., the results of the above measures except the last one are given from the various tests for Group 1 and for Group 2.

The first column in Table I., 'Highest No. Correct,' represents the memory-span measurements. The second, third, fourth, sixth, eighth, and tenth contain the measurements in the order described above. Test I. represents the tests made before the practice series and Test II. represents the tests made after the practice series. The averages opposite I. and II. are the averages, and the A.D.'s under them are the deviations, of the individual averages from the group averages. D_1 represents the differences between tests I. and II. for Group 1, while D_2 represents the corresponding differences for Group 2. A plus indicates the amount of gain in Test II. over Test I. A minus indicates a corresponding loss. D_1–D_2 the difference in gain or loss between groups 1 and 2. The A.D.'s in the vertical column after the rubrics, 'Error per Quantity,' 'Mems. Repr. per Quantity,' etc., are the deviations of the individual subjects in the test in question. The reason that the A.D.'s of 'The Error per Quantity' are so large is that the increasing the quantity of learning material beyond the memory-span increases the amount of error enormously. Thus, for example, when five letters were read the amount of error in the reproduction for most of the subjects was zero, but when a series of twenty letters were read, the amount of error for most of them was between sixteen and seventeen. This feature will become clearer when we shall discuss the influence of the quantity of learning material upon the reproduction.

From Table I. it will be seen that the value of the 28 gains of Group 1 in Test II. is destroyed by the counterbalance of 19 losses, by the fact that only two of them were more than twice as great as the A.D.'s of the averages from which they

Table I

Group	Test		Highest No. Correct	First Excess of ⅓ Error	First Excess of ½ Error	Error per Quantity	Aver. A.D.	Mems. Repr. per Quantity	Aver. A.D.	Size Largest Group Repr. per Quantity	Aver. A.D.	No. of Positions Correctly Repr. per Quantity
		Letters										
1	I	Ave.	6.75	8.12	9.12	7.48	4.10	6.53	1.37	3.25	1.35	4.47
		A.D.	1.31	.91	1.13	.52	.36	.25	.14	.47	.21	.60
	II	Ave.	6.75	7.13	8.88	9.23	4.78	7.30	1.36	3.66	1.31	4.67
		A.D.	.63	.56	.91	.51	.37	.59	.26	.48	.16	.47
		D_1	000	− .99	− .24	−1.75		+ .77		+ .41		+ .20
2	I	Ave.	8.00	9.67	10.50	7.47	4.99	7.44	1.25	3.63	1.50	4.76
		A.D.	.67	.78	1.20	.79	.32	.74	.13	.63	.19	.75
	II	Ave.	7.33	7.67	8.83	9.66	5.53	7.03	1.43	3.57	1.30	4.65
		A.D.	.44	1.67	2.16	.79	.52	.50	.22	.59	.16	1.02
		D_2	− .66	−2.00	−1.67	−2.19		− .41		− .06		− .11
		D_1–D_2	+ .66	+1.01	+1.43	+ .44		+1.18		+ .47		+ .31
		Numbers										
1	I	Ave.	8.00	9.38	10.75	7.47	4.59	8.56	1.39	4.82	1.34	5.75
		A.D.	1.50	1.87	1.75	.39	.78	1.13	.29	.84	.42	.81
	II	Ave.	8.25	9.75	10.75	5.76	4.18	8.01	1.60	4.97	1.42	6.27
		A.D.	1.31	1.44	1.25	1.60	.30	.73	.25	.65	.43	.82
		D_1	+ .25	+ .37	− .50	+1.71		− .45		+ .15		+ .52
2	I	Ave.	8.83	10.50	11.83	6.86	4.65	9.18	2.10	4.96	1.63	6.11
		A.D.	1.11	1.50	.89	.55	.57	.51	.35	.80	.11	1.04
	II	Ave.	9.00	11.00	12.17	5.38	4.08	4.48	1.55	4.83	1.62	6.31
		A.D.	1.67	1.67	1.83	.84	.52	.67	.24	.96	.13	.88
		D_2	+ .17	+ .50	+ .34	+ .48		− .70		− .13		+ .20
		D_1–D_2	+ .08	− .13	− .84	+1.23		+ .25		+ .28		+ .32
		Syllables										
1	I	Ave.	4.37	5.38	6.13	5.40	3.08	4.15	.74	1.68	.89	2.02
		A.D.	.81	1.13	1.09	.36	.68	.48	.11	.43	.36	.64
	II	Ave.	4.54	6.50	7.13	5.55	3.44	4.74	.88	1.96	.94	2.54
		A.D.	.71	.63	.88	.24	.52	.46	.14	.41	.19	.37
		D_1	+ .17	+1.12	+1.00	− .15		+ .59		+ .28		+ .52
2	I	Ave.	5.17	6.50	7.00	5.35	3.23	4.47	.71	1.92	1.03	2.32
		A.D.	.83	.67	.67	.89	.42	.54	.10	.48	.25	.55
	II	Ave.	4.66	6.33	6.83	5.56	8.29	4.54	.83	2.10	.89	2.59
		A.D.	1.83	1.67	1.17	2.80	.16	.47	.17	.42	.22	.67
		D_2	− .51	− .17	− .17	− .21		− .07		+ .18		+ .27
		D_1–D_2	+ .68	+1.29	+1.17	+ .06		+ .66		+ .10		+ .25
		Words										
1	I	Ave.	5.13	6.25	6.50	6.39	3.65	5.34	1.19	2.84	1.09	3.36
		A.D.	.91	.75	.75	.50	.54	.50	.28	.51	.24	.73
	II	Ave.	5.00	6.00	6.00	6.47	3.45	5.59	1.04	2.79	.95	3.73
		A.D.	.67	.78	.75	.44	.45	.51	.27	.46	.17	.52
		D_1	− .13	− .25	− .50	− .08		+ .25		− .05		+ .37

Table I—*Continued*

Words—*Concluded*

Group	Test		Highest No. Correct	First Excess of 1/3 Error	First Excess of 1/2 Error	Error per Quantity	Aver. A.D.	Mems. Repr. per Quantity	Aver. A.D.	Size Largest Group Repr. per Quantity	Aver. A.D.	No. of Positions Correctly Repr. per Quantity
2	I	Ave.	5.67	7.17	7.33	6.57	4.11	5.42	.91	2.93	1.01	3.69
		A.D.	.56	.89	1.11	.72	.31	.41	.27	.46	.13	.57
	II	Ave.	3.17	6.50	6.50	6.95	4.04	5.00	.80	2.53	1.00	3.27
		A.D.	.28	.67	.83	.45	.56	.33	.26	.42	.09	.58
		D_2	− .50	− .67	− .83	− .38		− .42		− .40		− .42
		D_1–D_2	+ .37	+ .42	+ .33	− .53		+ .67		+ .35		+ .79
		Vocabularies										
1	I	Ave.	4.64	4.71	4.71	4.35	2.38	3.27	1.02	3.45	.97	3.18
		A.D.	.23	.08	.08	.52	.10	1.16	.21	.96	.48	1.04
	II	Ave.	4.29	4.30	5.14	3.67	1.96	3.87	1.10	2.60	.99	3.57
		A.D.	1.40	1.11	1.63	.97	.29	1.00	.21	.94	.43	1.44
		D_1	− .35	− .41	+ .43	+ .68		+ .60		+ .15		+ .39
2	I	Ave.	3.80	5.00	5.20	4.17	2.38	3.05	.78	2.23	.86	2.42
		A.D.	1.00	1.20	1.04	1.49	.36	.54	.31	.52	.08	.77
	II	Ave.	3.40	4.80	5.60	4.06	2.38	3.58	.74	2.31	.82	2.89
		A.D.	.88	.64	1.12	.68	.44	.40	.11	.60	.11	.80
		D_2	− .40	− .20	+ .40	− .09		+ .53		+ .08		+ .47
		D_1–D_2	+ .75	+ .21	+ .03	+ .77		+ .07		+ .07		− .08
		Poetry										
1	I	Ave.	18.55	18.28	19.13	9.21	7.87	14.29	3.24	9.88	4.08	11.76
		A.D.	2.19	1.07	1.62	3.34	1.47	1.71	.67	2.06	1.11	2.68
	II	Ave.	10.32	14.86	16.30	12.72	6.43	13.11	2.93	8.02	2.60	10.35
		A.D.	4.40	2.45	3.19	3.07	1.02	1.99	.85	2.36	.86	5.73
		D_1	−8.23	−3.42	−2.83	−3.51		− .18		−1.86		−1.41
2	I	Ave.	18.60	20.80	20.80	7.66	7.01	11.75	3.38	16.00	3.10	14.22
		A.D.	2.28	2.24	2.24	1.60	1.15	1.65	.64	1.35	.40	1.83
	II	Ave.	14.20	17.20	17.60	12.25	8.17	14.06	3.52	9.35	2.84	11.33
		A.D.	1.52	1.44	1.92	2.37	1.86	1.36	.46	1.31	.32	2.21
		D_2	−4.40	−3.40	−3.20	−4.59		+2.31		−6.65		−2.89
		D_1–D_2	−4.83	+ .18	+ .37	+1.08		−2.49		+4.79		+1.48
		Prose										
1	I	Ave.	15.00	17.00	17.50	15.04	9.97	14.59	8.44	8.84	3.46	10.36
		A.D.	1.50	1.50	1.50	1.44	.66	1.55	.35	1.13	.86	1.52
	II	Ave.	17.00	20.00	20.50	13.07	8.09	16.43	3.86	10.70	3.50	12.92
		A.D.	3.50	2.50	2.15	2.15	.79	1.55	.34	1.99	.82	2.06
		D_1	+2.00	+3.00	+3.00	1.97		1.84		1.86		2.56
2	I	Ave.	16.66	16.33	21.33	14.70	1.24	14.90	3.08	9.40	3.48	10.48
		A.D.	1.88	2.33	3.33	2.56	.74	3.12	.47	1.75	.18	1.46
	II	Ave.	20.00	22.00	23.20	12.45	8.66	17.23	3.71	11.05	3.54	11.55
		A.D.	3.20	.00	1.92	3.23	1.52	2.05	.95	2.23	.54	1.68
		D_1	+3.34	+5.27	+1.87	+1.25		2.33		1.65		1.07
		D_1–D_2	−1.84	−2.67	+1.13	+ .72		+ .49		+ .21		+ .49

were computed, and by the 19 gains made by Group 2 in the same tests. When the difference in gain and loss between the groups is considered, we discover that Group 1 had thirty-nine gains over Group 2, and against these ten losses must be counted. This situation in itself apparently indicates that Group 1 increased its memory-span somewhat by the training in nonsense syllables, but, when the character of the gains and losses which make up this situation are considered, the facts of Table II. can only mean that the *special training had no effect in either increasing or decreasing the memory-span* for the various materials with which the tests were made. The important condition of this conclusion is that neither Group 1 nor Group 2 made any significant gains or losses in Test II. as compared with Test I.

Table II. shows the average performances of Groups 1 and 2 in the learning tests. *L* = Learning; *RL* = Relearning; ′ = minutes and ″ = seconds. The other symbols are the same as in Table I. The A.D.'s represent the deviation of the individuals of the group from the group average.

Table II. gives the measurements of the average performances of Groups 1 and 2 in the learning tests. Group 1 made three gains and three losses in the learning functions. Only one of the gains is greater than the A.D. of the average of either Test I. or Test II., and the other two are less. Group 2 made three gains and three losses in the learning functions, and all of them are less than the A.D.'s of the averages of either Test I. or Test II. This can only mean that there is clearly no *general learning ability that is improved by the special* training in nonsense syllables. Although no general learning ability is improved by the special training in nonsense syllables, it is still possible that this training might have improved some special learning function. In the test with nonsense syllables, Group 1 gained 6′ 47.3″ over Group 2. In the diagrams, Group 1 gained 5′ 12.9″ over Group 2. These gains are much greater than those gained in any other functions. When these gains are not considered with reference to the A.D.'s, they indicate that the special training did improve the ability to learn nonsense syllables and nonsense

Table II

	Test	Syllables ′	L″	A.D.″		RL″		A.D.″		Poetry ′	L″	A.D.″		RL″		A.D.″	
1	I	12	44.2	6	11.9	4	47.0	1	53.1	21	29.1	13	45.5	3	45.7	1	27.5
	II	7	34.7	4	5.6	3	5.6	1	10.4	19	16.4	10	49.0	3	29.4	1	32.4
	D_1	+ 5	95			+1	41.4			2	12.7			+0	16.3		
2	I	7	7.2	2	11.6	3	52.4	1	24.8	18	44.5	6	45.6	4	29.7	1	40.2
	II	8	46.5	3	32.6	3	29.0	0	52.9	18	49.9	6	6.8	4	52.4	2	32.8
	D_2	− 1	39.3			+0	23.4			− 0	5.4			−0	22.7		
	D_1–D_2	+ 6	47.3			+1	18.0			+ 2	18.1			+0	39.0		
		Diagrams								Prose							
1	I	17	18.2	6	10.3	5	57.5	3	12.4	23	22.7	8	51.8	7	35.5	5	24.0
	II	9	36.3	4	37.3	3	38.9	0	59.6	23	27.7	10	27.7	5	18.8	3	10.7
	D_1	+ 7	31.9			+2	18.6			− 5	05.0			+2	16.7		
2	I	12	11.6	3	8.4	5	49.7	0	50.4	29	48.3	13	54.1	5	38.5	1	0.6
	II	9	52.6	2	55.5	3	20.4	1	4.2	27	43.0	11	20.4	8	2.3	3	37.7
	D_2	+ 2	19.0			+2	29.3			+ 2	05.3			−2	23.8		
	D_1–D_2	+ 5	12.9			−0	10.7			− 2	10.3			+4	40.5		
		30 Vocabularies								40 Vocabularies							
1	I	18	34.8	5	10.7	7	5.6	4	53.3	23	59.0	9	34.9	7	11.8	4	17.9
	II	19	58.5	6	41.9	5	42.0	2	43.3	27	34.7	7	30.0	8	40.8	3	30.7
	D_1	− 1	18.7			+1	23.6			− 3	35.7			−1	29.0		
2	I	23	29.8	7	50.5	8	29.6	3	55.5	31	34.6	13	22.1	7	51.8	4	6.4
	II	18	38.3	6	37	6	34.5	1	42.8	35	55.6	15	19.4	6	36.2	1	46.1
	D_2	+ 4	50.5			+1	55.1			− 4	11.0			+4	15.6		
	D_1–D_2	− 6	9.2			+	31.5			+ 0	35.3	− 2	44.6				

visual diagrams. We must bear in mind, however, that Group 1 might gain more in a given test than Group 2, without the greater amount of improvement being due to special training.

In the relearning functions, Group 1 has five gains and one loss, but in all cases, at least one of the A.D.'s of the average from which these gains or losses were computed, is greater than the gain or loss in question. Group 2 has four gains and two losses. Only one of the gains is greater than the A.D. of the group averages from which it was computed. These facts undoubtedly mean not only that there is no general relearning ability which is improved by special training in learning nonsense syllables, but also that the efficiency of no special learning function is apparently affected by the special training.

The measurements given in Tables I. and II. may be supplemented by the measurement of the Pearson coefficient of correlation. If an individual's memory-span and his learning ability are allied, or if both are allied to some general ability, and if one of these factors is improved by a special training series, there ought to be a positive correlation between the performances in the training series and the performances in the test series, and the correlation between the training series and the second test series ought to be higher than that between the training series and the first test series.

Since there were seven measurements in the memory-span tests, it is possible that one of these is more reliable than another. The rank of the individuals in each test in a given measurement was correlated with their rank in the same test in each of the other measurements. For example, the memory-span for letters was correlated with the amount of error in letters, with the average size of the largest group reproduced in letters, and with the average number of positions correctly reproduced in letters. The 'First Excess of 1/3 Error' and 'The First Excess of 1/2 Error' were not considered in the correlation. Only the performances of Group 1 in Test I. were considered. Tables III. and III. *A* show the results of these computations. In Table IV. are shown the correlations

Table III

Showing the Correlation, R, between the Different Measures of the Performances in the Memory-Span Tests, Group 1, Test II.

$$R = 1 - \frac{6\Sigma g}{n - 1}$$

Tests

Measures / Correlation, R, of Memory-span with	Letters	Numbers	Syllables	Words	Vocabularies	Poetry	Prose	Ave.
Amount of error	.05	.48	.14	.48	.91	.52	.48	.44
No. of members	.24	.29	.53	.48	1.00	.22	.22	.43
Size of groups	.64	.53	.48	.48	.91	.48	.68	.60
Positions	.43	.48	.38	.22	.91	.31	.48	.46
Amount of error with no. of members	.14	.81	.34	.34	.91	.81	.34	.53
Size of groups	−.14	.53	.15	.71	1.00	.91	.71	.59
Positions	.24	.72	.24	.64	1.00	.81	.67	.62
No. of members with size of groups	.23	.48	.34	.43	.91	.72	.53	.52
Positions	.34	.62	.34	.14	.91	.81	.57	.53
Size of groups with positions	.22	.48	.81	.53	1.00	.91	.67	.66
Ave.	.24	.54	.38	.45	.95	.65	.54	

between each test in a given measure and the same test in other measures. In Table III. *A* the averages of these seven correlations between any two measures are shown. The general averages given at the bottom of Table III. *A* may be considered as giving the relative reliability of the five measures

Table III*A*

Showing the Average Correlations between the Different Memory-Span Measures as Computed from all the Tests, the Individual Correlations of Table III

	Memory-span	Amount of Error	No. of Members	Size of Groups	Position
Memory-span		.44	.43	.60	.46
Amount of error	.44		.53	.59	.62
No. of members	.43	.53		.52	.53
Size of groups	.60	.59	.52		.66
Positions	.46	.62	.53	.66	
Ave.	.48	.55	.50	.59	.57

that were correlated. If so, the Size-of-the-Groups measure is the most reliable, having a correlation of .59; and the Memory-span measure is the least reliable, having a correlation of .48. In the same way, the averages at the bottom of

Table III. may be considered as giving the relative reliability of the seven memory-span tests, and, if so, the test with vocabularies is the most reliable, having a correlation of .95; and the test with letters is the least reliable, having a correlation of .24.

Having determined the reliability of the different measurements, the next step was to measure the spread of improvement from the training series by the Pearson coefficient of correlation. The rank of the individuals in the training series was determined by the average learning time of all the syllable

TABLE IV

SHOWING THE CORRELATIONS, r, BETWEEN THE AVERAGE LEARNING ABILITY, A.L.A., IN THE PRACTICE SERIES WITH EACH OF THE TESTS BEFORE (I) AND AFTER (II) THE PRACTICE

$$r = \frac{\Sigma x \cdot y}{\sigma x \sigma y}$$

A. Memory-Span Tests, with Memory-Span Measure

	Letters		Numbers		Syllables		Words		Vocabs.		Poetry		Prose	
Test.............	I	II	I	II	I	II	I	II	I	II	I	II	I	II
A.L.A............	.69	.32	.65	.47	.23	.86	.52	.15	.66	.68	.44	.82	.30	.58

B. Memory-Span Tests, with Size of Group Measure

	Letters		Numbers		Syllables		Words		Vocabs.		Poetry		Prose	
Test.............	I	II	I	II	I	II	I	II	I	II	I	II	I	II
A.L.A............	.60	.65	.59	.60	.34	.38	.20	−.04	.51	.74	.35	.98	−.60	.66

C. Learning Tests

	12 Syllables		Diagrams A		30 Vocabs.		40 Vocabs.		Poetry		Prose			
Test.............	I	II	I	II	I	II	I	II	I	II	I	II		
A.L.A............	.91	.95	.26	−.47	.95	.77	.93	.81	.84	.77	.83	.94		

rows learned in that series. Their rank in the memory-span tests was determined both by the memory-span measure, since it was the principal measure of Ebert and Meumann, and the Size-of-the-Groups measure, since it had the greatest reliability. The rank of the individuals in the learning tests was determined by the learning time since it was the only measure used. Table IV. gives the results.

The facts of Table IV. *A* show a rather high positive corre-

lation between the average learning ability for nonsense syllables and the memory-span for each of the first series of the tests. These correlations are improved in the second series of tests in the case of nonsense syllables, vocabularies, poetry, and prose; and they are reduced for the other tests.

The correlations between the average learning ability and the memory-span tests, as determined by the Size-of-the-Groups measure, which are given in Table V. *B*, compare favorably with those of Table V. *A*. There are two cases, however,

TABLE V

THE RESULTS ARE IN MINUTES, SECONDS, AND PERCENTAGES

A. Ebert and Meumann

	(1)	(2)	D_1	Per Cent. 1
Learning	3′ 22″	2′ 1.6″	1′ 20.4″	59.7
Relearning	1′ 24.2″	1′ 6.0″	0′ 18.2″	77.9

B. Test Experiment

	(1)	(2)	D_1	Per Cent. 1
Learning	8′ 27.8″	5′ 14.7″	3′ 13.1″	63.7
Relearning	3′ 40.1″	2′ 34.9″	1′ 5.7″	74.8

where the relationship between Tests I. and II. changes direction, namely, in the tests with letters and numbers.

The correlations of the average learning ability in the training series with the first series of learning tests are on the whole higher than those with the memory-span tests, which indicates a closer relationship between these functions. These correlations are increased in the second series of learning tests for nonsense syllables and for poetry. The rank in the case of the first was determined by the twelve-syllable test series alone, since it had the same quantity as those in the test series. The most significant gain is made with this test, thus confirming the result of Table II., but in case of the diagrams, a negative correlation of .47 appears in the second test as against the positive correlation of .26 in the first test. This drop of .73 means that the individuals who did the worst in the training series did the best in the second test with the diagrams. But as measured by the learning time in Table II., more gain was made in this test than in any other one. If

the negative correlation has any significance, it certainly means that the improvement as measured by the time is not due to the training series but to the discovery of some special trick by which these nonsense visual diagrams were learned with ease. Where the measurements of Tables I. and II. contradict those of Table V., it would be hazardous to express an opinion in regard to the transfer of training. Where they do not contradict each other, it may be safe to do so. In the learning tests, there is only one such case that indicates a transfer effect from the training series, namely, the test with nonsense syllables. In the memory-span tests, there are two cases, if the gains of Group 1 over Group 2 are taken in an absolute sense without reference to the A.D.'s, namely the memory-span for nonsense syllables and the memory-span for vocabularies. But the only safe conclusion appears to be that improvement in learning nonsense syllables by one method improves the ability to learn nonsense syllables by another method. In any case, the spread of improvement from a special function is not general, but it is very specialized and affects only such other special functions that are very similar to the one specially trained. The result of the Test Experiment in regard to the transfer of training is that the findings of Ebert and Meumann and the conclusions based upon them, that there is a general memorial function and that special training in one function improves the memory in general, are not confirmed. Nor is their theory confirmed that transfer from one function to another is in proportion as the functions are allied. If the correlations in Table IV. are measures of the relations between the learning ability for nonsense syllables and the special functions experimented upon in the test series, they mean that the specially trained function is closely related to practically all of the functions measured in the test series, and that the relationship of this function to the learning functions is closer than to the memory-span functions. Either the learning and memory-span functions condition each other or both sorts are conditioned by some unknown general ability. Excepting the diagrams, the specially trained function has a relationship with all of the learning functions of the learning

tests above .75, but there is evidence of transfer to only one of these and evidence of interference with two of them. The probable meaning of this is that the conditions of improvement and of transfer are peculiar to the function especially trained and do not apply to any intrinsic relationship between the abilities of the various mental functions. What these peculiar factors are is difficult to say, but the writer has some grounds for believing that they are the special associative connections developed in the course of practice upon a given material, and in proportion as these peculiar associative connections can be applied to other materials, there is a transfer of improvement.

The next question is whether the results of Ebert and Meumann and the results of the test experiment are comparable. Did the subjects of the test experiment have as much to transfer as those of Ebert and Meumann: that is, did they improve as much? If they did not, the results of the test experiment could not be considered as modifying the results of Ebert and Meumann. We may compare the amounts of improvement attained in the two experiments by the following methods: 1. We may calculate the average learning time, (1) of the first two-syllable rows learned by each method in the training series, the same (2) for the last two by each method, and then take the ratio, per cent. 1, of the last to the first as the measure of the amount of improvement. 2. We may take the absolute difference, D_1, between these two learning times as the measure of the amount of improvement. 3. We may take the ratio, per cent. 3, of the average learning time of the last two days, (3) of the practice period to that, (4) of the first two days of the practice period. 4. We may take the absolute difference, D_3, between these two as the measure of the amount of improvement. 5. We may take all of these together. The writer has calculated all these measures from the data of the two experiments in question. The results are given in Table V. Table V. *A* shows the group averages for Ebert and Meumann's Subjects while Table V. *B* shows them for Group 1 of the test experiment.

It will be noticed that the amount of improvement as measured by the ratio of the last two of each method to the

first two of each method in the practice period of Ebert and Meumann's subjects reduced their learning time 4 per cent. more than those in the Test Experiment, 59.7 as against 68.7; but as measured by the ratio of the average learning time of the last two days of practice to the first two days of practice, they improved 10.9 per cent. less than those in the test experiment, 60.4 as against 49.5. The reason for this difference is that Ebert and Meumann's subjects learned their series as regards method in a fixed order, so that each subject covered the four methods used by them every two days. This was not the case in the test experiment in which three or four days were often required to learn two series by each method. The ratios of the last two days of the practice period to that of the first two days should, therefore, be regarded as more reliable, and, according to this measure, the subjects of the test experiment improved more than did those of Ebert and Meumann in their first practice period. If the absolute difference between the learning times between the first two by

TABLE VI

A. Ebert and Meumann

	(4)	(5)	D_3	Per Cent. 3
Learning	3′ 50.3″	2′ 1.6″	1′ 20.4″	60.4
Relearning	1′ 19.2″	1′ 6″	0′ 13.2″	82.0

B. Test Experiment

Learning	10′ 17.8″	4′ 47.9″	5′ 29.9″	49.5
Relearning	3′ 55.0″	2′ 32.0″	1′ 23.0″	66.8

each method and the last two by each method, or if the difference between the learning times of the first two days of practice and the last two days of practice be regarded as the measure of the amount of improvement, then the subjects of the test experiment improved more than twice those of Ebert and Meumann. If the quantity of material learned is important for the question at hand, Ebert and Meumann's subjects each learned thirty-two series while each of those in the test experiment learned forty-two series in the average. And with respect to time, it appears that the subjects of the test experi-

ment spent more than twice as much time in practice as those of Ebert and Meumann. The learning times for the latter were obtained by multiplying the number of repetitions by twelve seconds, the duration of one rotation of the drum from which the series were read. It appears, however, that Ebert and Meumann stopped their drum while the learner made a trial reproduction, and then started it again from the syllable on which he first halted. Such a repetition would require more than twelve seconds, thus causing the above method of calculating the learning times to be too low. However, the conclusion from the discussion is that with regard to the amount of improvement and the condition of its transfer, the test experiment is fully comparable with the one of Ebert and Meumann. Not only did its subjects spend more time in practice and learn a greater quantity of material, but they also attained a greater amount of improvement both in learning and relearning than did the subjects of Ebert and Meumann in their first practice period. Other things being equal, the subjects of the test experiment should have shown more spread of improvement than those of Ebert and Meumann, but the former, unlike the latter, instead of showing that improvement n one function raises the efficiency of every other function to a large extent, showed that it principally raises the efficiency of the one specially trained and that of other special functions to a small extent, provided they are similar.

Just what the explanation of Ebert and Meumann's results must be in the light of this conclusion is difficult to say. But certain experiments of Swift, Schuyler, Thorndike, and Book have shown that improvement in relearning after several years of no practice is much more rapid than in the original learning. In view of the extraordinary amount of memorizing done in the German gymnasium, I believe that the improvement of Ebert and Meumann's subjects in their tests must be interpreted as a rapid re-acquisition of their former memorial skill learned in the gymnasium.

The second problem discussed by Ebert and Meumann was the economy of learning. In this connection, the series

of nonsense syllables of the practice period were learned by four methods as follows: *A*. This was the whole method. The series were repeated without pause from beginning to end until the first correct reproduction was reached. *B*. The part method, in which the subject first learned the first half of the series, then the second half, and finally the two halves together. *C*. The same as the *A* method, except that the subject paused one second at the end of the fourth syllable and one second at the end of the eighth syllable. *D*. The same as the *A* method except that the subject paused two seconds at the end of the sixth syllable. The essential difference between these two methods is the size of the units in which the respective series are learned. An *A* series is learned as a unit of twelve. A *B* series is first learned in units of six, and then the two units are learned together. A *C* series is learned together from the beginning but in units of four. And a *D* series is learned together from the beginning but in units of six. These four learning methods were also followed in the test experiment, but instead of the order of succession being varied in a fixed way, so that each subject learned an equal number by each, as with Ebert and Meumann, the order was varied irregularly so that in some cases a subject learned twice as many by one method as by another. The purpose of this was to discover whether the amount of practice would be a factor in determining the most economical method of learning. The group averages on the economy of learning, as the writer has calculated them from the data of both experiments, are given in Tables VII. and VIII.

TABLE VII

(EBERT AND MEUMANN)

Method	*A*		*B*		*C*		*D*	
S's.	L.	Rl.	L.	Rl.	L.	Rl.	L.	Rl.
Ave's	10.5	4.4	9.2	4.7	7.9	4.1	9.4	4.4
L. + Rl. of ave's		14.9		13.9		12.0		13.9
Rl. after 24 hrs.		56.9		47.5		49.7		63.4

Commenting upon these results, Ebert and Meumann say that the *C* and *D* methods are the quickest in learning and

produce a retention of average security. The *D* method is relatively quick in learning but uncertain in retention. The *A* method is the slowest in learning but has the greatest security in retention and certainty in reproduction. The *C* method is by far the best in learning. The *A* method is equally the best in retention, and, as regards the speed of learning, the *B* method lies nearer to the *C* and *D* methods than to the *A* method. In general these statements are confirmed in the test experiment. The *C* method is the quickest and the *A* method is the slowest in learning. As measured by the Saving

TABLE VIII

(TEST EXPERIMENT)

Times in Minutes and Seconds

Method	*A*				*B*				*C*				*D*			
S's.	L.		Rl.		L.		Rl.		L.		Rl.		L.		Rl.	
Ave's.	7	15.1	3	7.5	6	18.8	3	19.	5	49.9	2	55.2	6	48.5	2	29.4
L. + Rl.			10	22.6			9	37.8			8	45.1			9	17.1
Per cent. saved in Rl. after 24 hrs.				59.				48.3				48.3				51.7

Method, the *B* method has the poorest retention, but the *D*, and not the *A* method, has the best retention. With respect to the speed of learning, the order from the slowest to the quickest of the different methods in both experiments is the same, namely, *A*, *D*, *B*, *C*. In relearning, the order for Ebert and Meumann is *B*, *D*, *A*, *C*; and in the test experiment, it is *B*, *A*, *C*, *D*.

An important question with reference to the most economical method of learning is how to measure economy in learning. Upon this matter, Meumann says: "That method of learning is most economical which secures a particular memorial effect or attains a particular memorial purpose in the shortest time, with the least number of repetitions, and with the minimum degree of fatigue; and this method may be regarded as the most economical only with reference to this memorial effect and this memorial purpose. Of these three determinations, the learning time measures the economy of time; while the economy of energy is measured by the number

of repetitions, and in less precise form, by the amount of fatigue. No other accurate means of measurement is at our disposal.[1] Meumann puts much emphasis upon the particular memorial effect desired. One method may be most economical for immediate retention, a second for temporary retention, a third for permanent retention. The quality and quantity of the material may also determine the most economical method of learning. And the character of the associations as well as the character of the reproduction, its fidelity, rapidity, and completeness must also be considered. With all of these factors, the economy of learning becomes an extremely complex problem, and much experimentation will be needed before all of these factors can be evaluated. But it is difficult to see why the time element would not be sufficient measure of the economy of anyone of these factors, especially if economy is to be measured in terms of efficiency. For example, if one method is quickest in learning and another is the quickest in relearning, and we wish to know which is the quickest in respect to both, we simply consider that here are two pieces of work to be done, learning and relearning, and that method by which both may be done in the least time is the most economical.

In Table VII. the average number of readings required by Ebert and Meumann's subjects for both learning and relearning by each method, have been added. In Table VIII., the corresponding times appear for the subjects of the test experiment. According to this measure, the *C* method is the most economical in both experiments, a fact contrary to the conclusions in the investigation of Steffens[2] and Pentschew[3], both of whom made out strong cases for the *A*, or whole, method.

In the test experiment an effort was made to evaluate two factors that might influence the most economical method of learning, namely, practice and the memory-span. The subjects did not learn an equal number by each method. This afforded an opportunity for learning the effect of practice upon the economy of learning. The effect in question was

[1] 'Psychology of Learning,' p. 373.

[2] *Zeit. f. Psychol.* **22**, pp. 321, 465.

[3] *Arch. f. d. ges. Psychol.* 1903, **1**, p. 417.

determined by calculating the Pearson coefficient of correlation. This was a difficult matter, but space permits only to say that it resulted in +.14, a figure indicating some relationship but not enough to be of importance.

To determine the correlation between the memory-span and the size of the unit into which the different learning methods naturally divided the syllable rows, the average memory-span of each subject in all of the memory-span tests, *i. e.*, those before practice were calculated. Each of these averages was then arrayed with the size of the unit in which the subject learned the quickest as in Table IX.

TABLE IX

S's.	Ave. Memory-span	Most Economical Unit
Nl.	10.57	6
Cl.	10.30	4
Sy.	10	12
Gr.	9.85	12
Jn.	7.95	12
Bv.	7.20	6
My.	6.93	4
Bh.	6.76	4
Ave.	8.67	7.5

$r = .46$.

The Pearson coefficient of correlation as calculated from these averages is +.46, which shows that the size of the units in which the subjects learned is a much more important factor in the economy of learning than the amount of practice, and that it is a factor to be reckoned with in connection with this problem. The importance of this factor is still more evident when we compare the memory-span of Group 1 in nonsense syllables with the size of the most economical learning unit, as in Table X.

TABLE X

Method	Ave. L. T.	Size of Unit in Each Method	Ave. Mem. Span of Group 1 in Syllables
A	7′ 15.1″	12	4.37 Test I
B	6′ 48.5″	6	4.54 Test II
D	6′ 18.8″	6	
C	5′ 49.9″	4	Ave. 4.45
Ave.	6′ 30.6″	7	

$r = .91$.

From Table X. it will be seen that the most economical unit, 4, lies very near the average memory-span of Group 1, 4.45. The units of the *B* and *D* methods, which are 6, lie 1.15 farther away from the group memory-span. The unit of the *A* method, which is the slowest of all, lies 7.15 away from the group memory-span. The relationship between the quickest learning method of Group 1 and its most economical learning unit is beautifully expressed by the Pearson coefficient, which is +.91. The correlation between the optimum method and the optimum unit of the group is thus .91, while that corresponding to the individuals is .46. It is probable that the latter would approach the former if a greater number of individuals were compared, and, if so, we must consider the size of the unit in which an individual learns as the principal determinant of his optimum method. The size of this unit is conditioned by the learner's memory-span. Under these conditions, we could not advise the whole method as such, or the part method as such, or the modified whole method as such, as the optimum method, but rather that an individual should learn in such units that lie near his memory-span for the material in question.

It is possible, however, that the character of the associations may have something to do with the optimum method. It will be noticed from Table X. that the learning time for the *B* method is shorter than for the *D* method, but its relearning time is much greater, so that the per cent. saved in the *D* method is 63.4 as against 47.5 in the *B* method. Considering both the learning and the relearning times, the *D* method is more economical than the *B* method. In the *B* method, the halves of the series were first learned separately and then together, while in the *D* method the halves were learned together from the beginning. My. and Bh. reported that learning the halves of the *B* method together after having learned them separately required a new organization of the entire series. They could not put the two halves together mechanically, but had to break up the associations formed with the halves and build new ones, thus causing interference between the two sets of associations. Such a condition would

easily explain the small saving in the *B* method, for the interference would naturally weaken both sets of associations. On the same grounds, the great saving in the *D* method may be explained as being due to the absence of such conflicting associations. But the difference in the saving between the *A* and *C* methods is better explained as due to the difference in imprinting. So far as this experiment goes, the advantage of the *C* and *D* methods is due to the fact that they do not form conflicting associations but yet adjust the material to the learner's memory-span. This would suggest that a learner should generally follow a modified whole method with pauses in between subdivisions the size of which is determined by the memory-span for the material in question.

The third problem which comes up in this experiment but which was not considered by Ebert and Meumann is the influence of quantity of learning material upon the memory-span. The memory-span tests varied so widely in quantity that they afforded an excellent opportunity for studying the effect of the quantity of the stimulus upon the quality of the immediate reproduction. A study of this sort is not only interesting for its own sake, but also for the light it throws upon the problem of the economy of learning.

The quality of the reproduction was measured in four ways: (1) the average amount of error; (2) the average number of members reproduced; (3) the average size of the largest groups reproduced; and (4) the character of the position reproduced. The nature of the first three measures was explained in the first part of this report, but a little more should be said about the positions. These were divided into three classes: (1) those that were correctly reproduced with respect to their relative positions to either end of the series; (2) those that were misplaced; and (3) those that were omitted.

The measurements of the subjects' performances in a given quantity in a given test were all placed in vertical columns, and then the averages of the measurements added vertically for both Tests I. and II. irrespective of the group divisions were calculated in order to show the effect of the quantity of the stimulus upon the amount of error. The A.D.'s of the

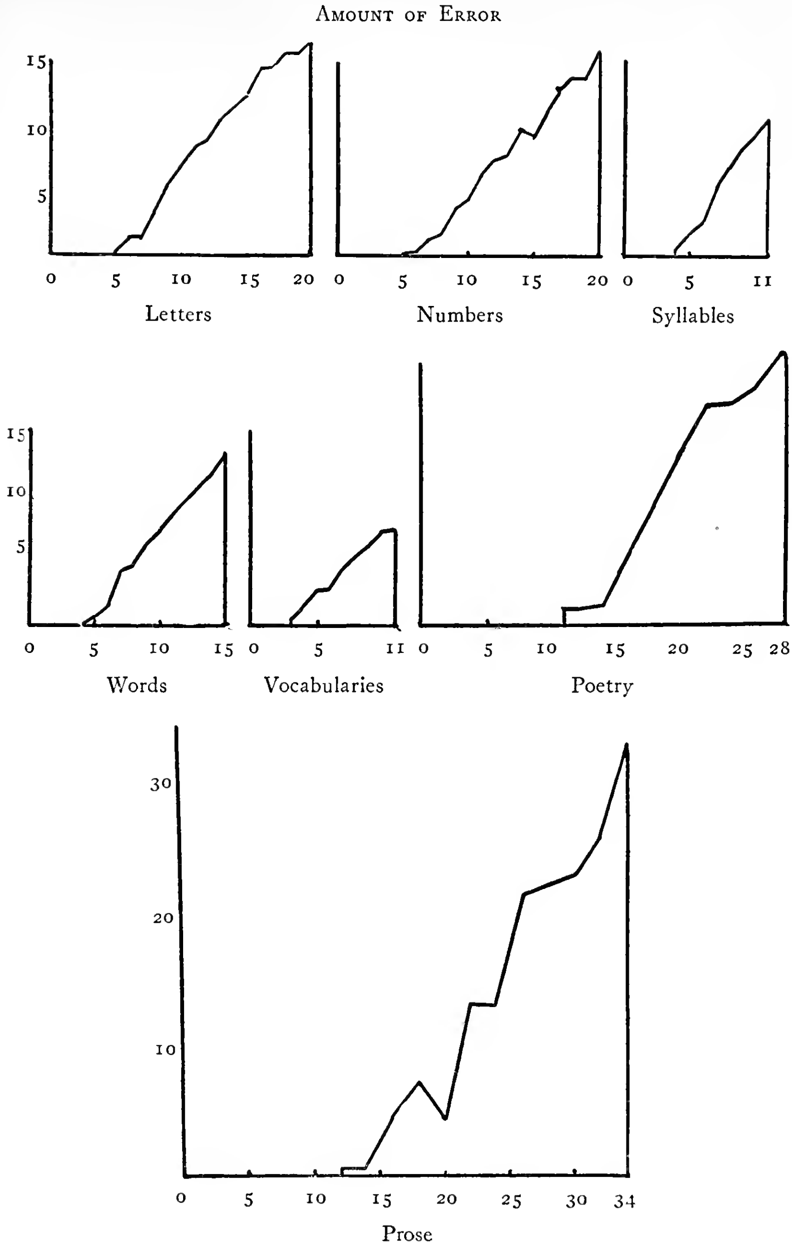


FIG. 3. Curves showing effect of quantity upon amount of error.
Ordinates = Ave. no. errors.
Abscissas = Ave. no. members in series.

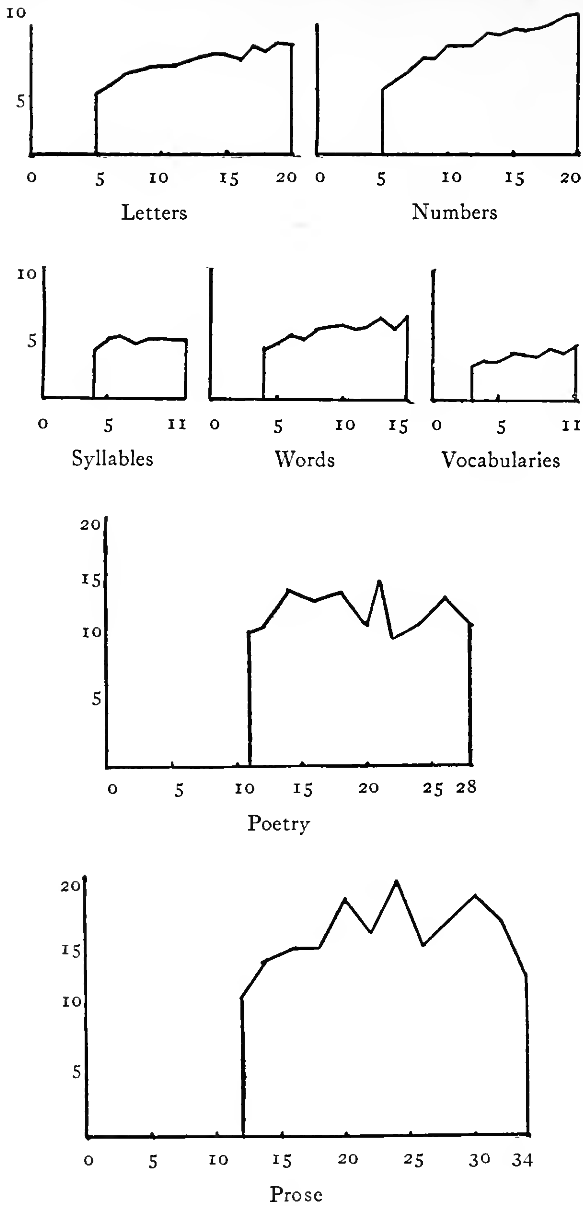


FIG. 4. Curves showing effect of quantity on number of members reproduced.
Ordinates = Ave. no. of members reproduced.
Abscissas = Ave. no. of members in series.

averages were also calculated. The averages in the tests with letters, numbers, and syllables are each based upon sixty measurements; and in those with vocabularies, words, poetry, and prose they are each based upon thirty measurements. The results with reference to the amount of error are presented graphically in Fig. 3. It will be noticed that the amount of error increases directly as the quantity of stimulus.

The results with reference to the number of members are presented graphically in Fig. 4. Perhaps the effect of quantity upon the number of members could be expressed by saying that increasing the stimulus beyond a quantum a little larger than the memory-span increases the number of members reproduced by a fraction whose value approaches zero. In fact, it is not very incorrect to say that beyond such a point the number of members reproduced is not increased. For example, in the case of syllables increasing the number of members beyond six does not increase the size of the reproduction beyond the average memory-span, 4.45. In case of prose, the number of members is not increased materially for sentences having more than 20 words, the average memory-span being a fraction over seventeen.

Fig. 5. shows the effect of quantity upon the size of the groups correctly reproduced. It will be noticed that increasing the quantity of the stimulus beyond the memory-span decreases the size of the groups correctly reproduced. This is more noticeably true of meaningful material than of nonsense material. For example, when a sentence of twenty words is read, the average size of the largest groups reproduced was 14.43 words. When a sentence of thirty-four words was read the average size of the largest group was only 4.75. When we consider that the ability to comprehend long spoken or written sentences depends not upon the ability to get a word here and there, but upon our ability to get a connection of all the words, we see the pedagogical importance of the above relationship between quantity and an immediate reproduction that is held responsible for order. This is all the more significant because most things that are read and heard are given only one repetition.

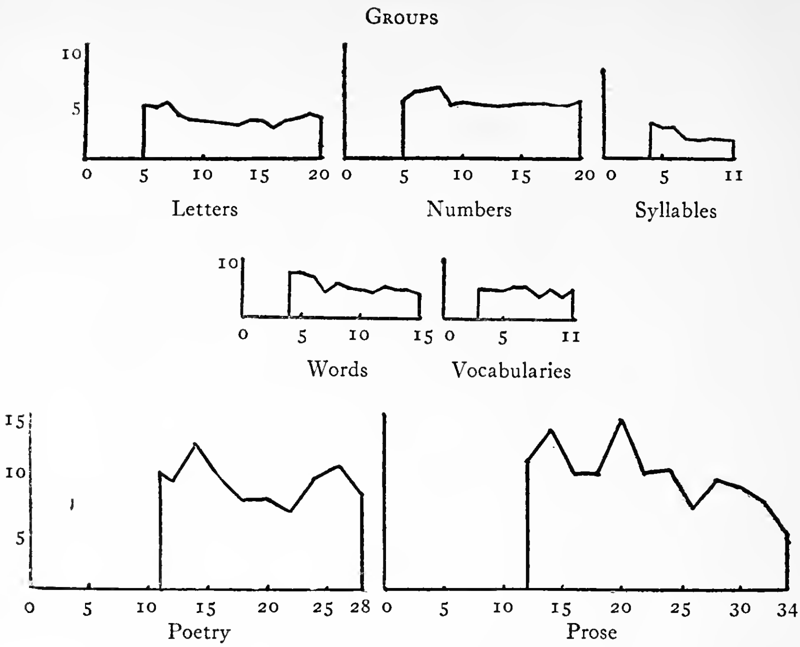


FIG. 5. Curves showing effect of quantity on size of groups.
Ordinates = Ave. size of groups reproduced.
Abscissas = No. of members in series.

The effect of quantity upon the character of the positions reproduced and omitted is shown in Fig. 6 which depicts those from the test with disconnected words. The curves here are typical of all the tests. They may be described verbally as follows: If in coördinates, the frequency of positions reproduced is plotted along the ordinates and the successive positions are placed along the abscissa line, the distributions of the positions are as follows: First, for those correctly reproduced: When the stimulus is less than the memory-span, their distribution is a rectangle. As the stimulus increases beyond this point, the roof of the rectangle begins to fall, first in the middle and then along the entire line, and, as it does this, the sides of the rectangle become a little shorter, but yet remain quite high even for large quantities, the side standing for the end of the series being higher than that for the first of the series. The roof of the rectangle thus tends to form a curve

which has the reverse shape of the probability curve. Second, the positions omitted: The curve for this omission is just the reverse of the one described. The distribution for quantities

Positions. Words.

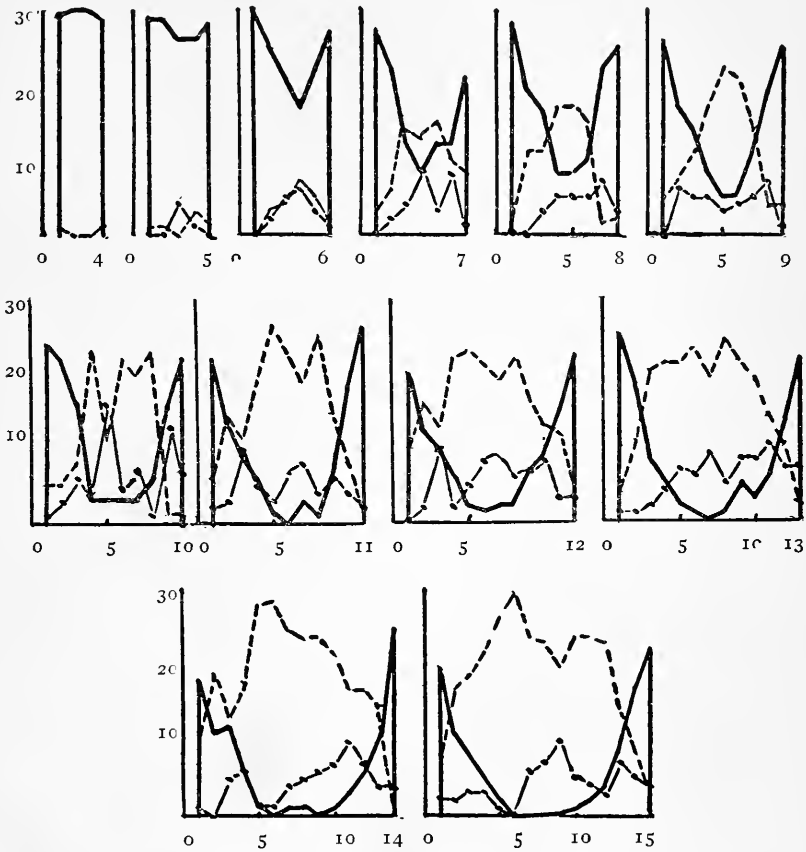


Fig. 6. Curves showing effect of quantity on distribution of positions correctly reproduced.

Continuous Lines = positions correctly placed.
Dots and dashes = positions misplaced.
Dashes only = positions omitted.
Ordinates = frequencies of reproductions.
Abscissas = order of positions in each series.

less than the memory-span is a zero abscissa line, and, as the quantity increases, this line forms a little lump in the middle

which grows larger and larger as the stimulus is increased, until the high point embraces all the positions in its ordinate; and from this, the line gradually falls toward the base at each end. Third, the misplaced positions: The distribution of these is irregular, but their tendency is to follow omissions.

The essential significance of these conclusions on the effect of quantity upon immediate reproduction is that the mind in learning a quantity of material must proceed by steps just as the body does in covering a quantity of space. The extent of that step is fixed just as much as the bodily step, and just as the body falls flat when it tries to take too large a step, so the mind literally falls when it tries, in one act, to comprehend a material that is much beyond its memory-span. It simply does not get it, and, in fact, it gets less than if it proceeded by small steps.

The significance of this for the economy of learning is evident. It means that the mind can learn only by parts, no matter what the method of reading is. If the whole method is followed, it means that different parts come above the memorial threshold at different times. If the part method is followed, fewer readings will be required to bring the part above this threshold. However, a correct reproduction must reproduce not only the members but also the order of the series. We have seen that learning in quantities greater than the memory-span reduces the learner's capacity for this kind of reproduction. Is it not then a saving of energy and time to adjust the optimum learning unit to the learner's memory-span? It appears that the above facts are a fatal blow to the much puffed-up whole method as such, and that the part method is much more in agreement with psychological laws. If so, the economy of learning as regards method is how to avoid conflicting associations.

THE INFLUENCE OF COLOR ON APPARENT WEIGHT. A PRELIMINARY STUDY

BY J. E. De CAMP

Stanford University

Contents

I. Statement of the Problem

The apparent weight of an object is a function of its size. Two objects of like material, appearance, etc., and having the same weight, but of different size, will appear unequal in weight not only to the unsuspecting subject, but as well, though to a less marked extent, to the experienced individual who is thoroughly familiar with the nature of the illusion involved. A second psychological factor influencing the weight of an object has been pointed out by Seashore.[1] This factor is the material out of which the object is thought to be constructed. The apparent weight of an object is influenced by many other factors than its size and material. It depends upon such factors as the neuro-muscular adjustment brought about by lifting one or more weights, the time elapsing between this adjustment and the lifting of the weight to be judged, the velocity with which the weight is raised from its support, the time occupied in the lifting process, whether it

[1] Univ. of Iowa Studies in Psychol., 1899, 2, 36–46.

is lifted before or after the weight with which it is to be compared, bodily conditions such as fatigue, etc.[1]

The question arose in the mind of the writer whether the color of an object might not influence its apparent weight. Would objects alike in all respects save that of color appear equal in weight? Is color in any way a determinant of the apparent[2] weight of an object? If the color of an object does influence its apparent weight we should like to know how the latter varies with a variation in hue, tint, and chroma. This paper and the experimental work herein reported will serve only as an approach to an answer to these questions. Its main object is to suggest, with experimental evidence for the suggestion, an answer to the particular question: Is color in any way a determinant of the apparent weight of an object? In other words, is there a color-weight illusion just as there is a size-weight illusion and a material-weight illusion?

II. Literature

So far as the writer is aware no experimenter has given a careful consideration to the problem with which we are now concerned. Seashore,[3] in connection with his experiments on the material-weight illusion, made a few tests on children to determine the presence or absence of a color-weight illusion. His method was that of selecting from a graded series a block thought to be equal to a comparison block. Only black and white were tested. His results indicated the absence of any color-weight illusion. It is obvious that the increment used

[1] Compare in this connection C. E. Seashore, 'Psychology in Daily Life,' N. Y., 1913, 175 ff.; G. E. Müller and F. Schumann, *Arch. f. d. gesamm. Physiol.*, 1899, 45, 37–112; L. J. Martin and G. E. Müller, 'Zur Analyse der Unterschiedsempfindlichkeit,' Leipsig, 1899.

[2] *Psychological* weight might be substituted for *apparent* weight. The qualification is necessary to distinguish the apparent weight when lifted by an experiencing person from the physical weight of the object. I have often noticed that the beginner in psychology not infrequently has difficulty in immediately realizing that the psychological weight of an object is just as real as its physical weight. The fact that reference is always made to a physical standard leads to the use of the term illusion. The pound of lead is really psychologically heavier than the pound of feathers. Scripture's law or formula (*Science*, N. S., 1897, 5, 227), of the size-weight illusion is only a means of passing from one real to another.

[3] Univ. of Iowa Studies in Psychol., 1899, 2, 45.

(5 grams) in his standard series of blocks was too large to separate out any factor whose influence is very slight. A second approach is some work done by E. Bullough[1] on 'The Apparent Heaviness of Colors,' which is, as the writer of the article points out, a contribution to the æsthetics of color. He is concerned in particular with the æsthetic effect of placing one color above another. It may be said that he attempts to investigate scientifically the rule "that dark colors should stand below light ones."[2] His method is that of choices or preferences. It is concluded that dark colors are heavier than light ones. Accoıding to him "it is the luminosity which forms the principal factor in the preferences,"[3] *i. e.*, the relative luminosity of a color determines whether it will be preferred above or below another color. Seemingly the empathic sense in which the term weight is here used differs quite markedly from that in which we use the term weight in speaking of a lifted object, but possibly the significance of the term in the two cases is quite the same. Our ignorance of the factors involved in the æsthetic arrangement of two or more colors at present precludes a definite answer to this question.

There are various articles dealing with the many factors influencing the apparent weight of a lifted object, such as the size of the object, the rate at which it is lifted, etc., but since we have endeavored to make such factors constants in our tests they need not claim our attention here, and it is not necessary to refer to the literature dealing with them.

III. Experimental Work

Section I

The experimental work reported in this section was carried on by the writer and two of his advanced students[4] in

[1] *British J. of Psychol.*, 1907, 2, 111–152.

[2] *Loc. cit.*, 113.

[3] *Loc. cit.*, 133; *cf.* also p. 152 where he says: "Since the saturated colors are not all of the same luminosity, but yellow and green are of higher luminosity than blue and violet, yellow and green are also of apparently lighter weight than blue or violet when seen singly."

[4] Misses Irene Hollenbeck and Laura Stephan. Sixty-five per cent. of the tests were made by the writer.

the psychological laboratory of the University of California during the second semester of the year 1915–1916. Of the sixty-three subjects tested nineteen were men and forty-four were women. The subjects were drawn from Professor Stratton's class in applied psychology. Seven of the subjects were seniors, thirty-five were juniors, nineteen were sophomores, and two were specials. The test was always given in the forenoon some time between 8 A. M. and 12 M. Due to a cloudy or foggy morning there were a few instances where the light conditions were slightly below normal (ordinary diffuse daylight). The test usually occupied about forty-five minutes, and was given to each subject individually. Each test fell into two main parts: for convenience we will call the first part Part 1, and the second part Part 2. Part 1 always preceded Part 2.

A. Apparatus.—The apparatus used was twenty-five cubical ($2\frac{1}{2} \times 2\frac{1}{2} \times 2\frac{1}{2}$ in.) soft pine blocks. These blocks were smoothed and made the desired weight by boring out part of the wood from the bottom of the block. Nine of the twenty-five were used in Part 1, all were used in Part 2. The nine blocks for Part 1 were all made exactly the same weight, seventy-six grams. Each one of the nine was covered with colored paper by glueing the paper evenly and closely to five sides of the block. The bottom was left uncovered. In this way the nine blocks were alike in size and weight; in fact they were alike in every respect save that of color. Each of the nine differed from every other one in respect to color. The following colors were represented: red, orange, yellow, green, blue, violet, purple,[1] black, and white. It will be seen later that the present work does not demand a refined photometric method such as is described by Sir Wm. Abney,[2] or like that given by Yerkes and Watson.[3] An attempt was made to obtain the approximate luminosity value of each color by the method suggested by Rood.[4] The following values were received for the luminosity of the different

[1] Probably more correctly designated by the commercial term *cerise*.

[2] 'Researches in Color Vision,' London, 1908.

[3] 'Behavior Monographs,' 1911, No. 2.

[4] *Amer. J. of Sci. and Arts*, 1878, **15**, February.

colors: red, 5.2; orange, 27.2; yellow, 75.7; green, 44.3; blue, 14.7; violet, 8.5; purple, 6.6; black, 0; white, 100; gray, 16.9. The figures given represent the percentage of white contained in the colored discs. No pretense is made to absolute accuracy. There can be little doubt, however, that the relative luminosity of the different colors is correctly shown.

The sixteen additional blocks used in Part 2 were of the same material and size as those used in Part 1. These were covered on five sides with gray paper (No. 16[1] of Hering's series). They all differed in weight. The sixteen made up a graded series from fifty-five to one hundred grams, the increment in each case being three grams. The set of sixteen blocks were alike, then, in every respect save that of weight.

Additional apparatus consisted of (*a*) a stand supporting a horizontal rod 16 cm. above the base of the stand, (*b*) a second stand supporting a black cardboard screen, and (*c*) a large piece of soft black velvet cloth. The horizontal rod mentioned under (*a*) was wrapped with black cloth.

B. Method: Part 1.—The general method used in Part 1 was that of Paired Comparisons. This method is so well known that it needs no discussion here.[2] The ranks and files of the table constructed for the thirty-six (*i. e.*, between nine objects; $\frac{9 \times 8}{2} = 36\Big)$ possible comparisons were headed thus:

	R	O	Y	G	Bl	V	P	Bk
O	1							
Y	2	3						
G		4	etc.					
Bl								
V								
P								
Bk								
W								

The observer[3] was seated at one side of a fairly large table, the experimenter at the other. The table was covered with the large piece of black velvet cloth. The stand supporting

[1] W. Brown has shown (Psychol. Rev., 1915, 22, 520 f.) that this particular gray is the seventeenth (instead of the sixteenth as numbered) in Hering's set.

[2] See E. B. Titchener, 'Experimental Psychology,' Vol. 1, Pt. 1, 92 f.

[3] Throughout this paper the terms *subject* and *observer* are used interchangeably.

the horizontal rod was placed near the observer. The purpose of the rod was to give uniformity to the height of each lift made by the observer. The height of the rod above the table was 16 cm. The observer was definitely shown how the lift was to be made. The elbow was to be used as a fulcrum; no other part of the arm or hand was to be bent. The subject[1] was told to have as nearly as possible a uniform method of grasping the blocks. The blocks were placed in position by the experimenter for the subject. Only two blocks, those being compared, were allowed on the table at the same time; the others were kept in the table drawers on the experimenter's side of the table. The cardboard screen was placed in such a way that the protocol of the experimenter was concealed from the subject's view.

The first comparison was made between red and orange, the second between red and yellow, the third between orange and yellow, etc. Care was taken that in the eight comparisons belonging to any color the block representing the color in question should be presented four times before and four times after the blocks with which it was compared. This served to eliminate any temporal error. All blocks were lifted from the same position, in that way avoiding any positional error. Before the test began the observer was impressed with the following four points:

1. The experiment is a comparison of weights.

2. The weights are all different; no two have the same weight.[2]

[1] *Vide* note 3, p. 351.

[2] It might be thought that the proper procedure would be to omit any reference to the weight of the blocks. In such a case should the observer be presented in the first few trials of the experiment with blocks whose color has little relative influence upon their weight he would very likely get the impression that the blocks were all equal in weight. The influence of this idea would invalidate the results obtained so far as the factor under investigation is concerned. Suppose that the different colors have little relative influence upon the apparent weight of objects, the introduction of this common factor will greatly aid in the determination of the little and will not materially affect its validity. The suggestion will aid in fixing in the mind of the observer the idea that the aim of the experiment is a comparison of weights and will tend to exclude from him any knowledge of the factor under consideration. This is highly desirable since the untrained observer can not avoid distorting actual occurrences; his results are biassed by what he thinks the results ought to be under such conditions.

3. The second weight in each comparison is to be judged in terms of the first.

4. The eyes are to be kept fixed upon the block while it is being lifted.

After these instructions had been given a few preliminary trials were usually made with some of the blocks with the graded gray series for the purpose of (1) familiarizing the observer with the experimenter's method, and (2) setting up a kind of muscular adjustment in each individual case for the perception of differences in weight. At first fairly large differences were given, then smaller ones. In these preliminary tests the subject was informed whether his judgments were right or wrong.

Part 1 was then conducted in accordance with the method of paired comparisons. The observer was in no case informed whether his judgment in any particular instance was correct. Sometimes when the observer showed too much anxiety about the correctness of his judgments the experimenter made such casual remarks as, "You are doing very well," or, "You are getting along very nicely." Immediately upon each individual judgment the experimenter entered the judgment in the table previously prepared for that purpose. Unless the observer showed signs of fatigue the test was carried directly through. This part of the test required approximately fifteen minutes. In case of fatigue, and this very seldom happened, the observer was allowed a brief respite.

Method: Part 2.—In this part the complete set of twenty-five blocks was used. Only two blocks were on the table at a time. Beginning with the red block and taking the colored blocks in spectral order each one served as a standard with which a number of the sixteen gray blocks were compared. In each case the colored block was lifted first, then the gray block. The gray blocks were taken in the order of their weight, *i. e.*, 55, 58, 61, ..., 100. Before beginning this test the observer was told that in this case some of the blocks compared might be just alike in weight, in this way indicating to the observer that in some of the comparisons 'the same' might occur as a correct judgment as well as 'lighter' or

'heavier' in other comparisons. The observer was not informed that the gray blocks constituted a graded series.[1] Record was made by the experimenter of the subject's judgments. In the first few experiments the comparison of any colored block with the graded series of grays was stopped upon the subject's first judgment of 'heavier.' In the later tests the comparison was stopped only when one judgment of 'heavier' was followed in the next comparison by a second judgment of 'heavier.' It was assumed that a continuation of the series would lead to a judgment of the remaining heavier weights as 'heavier.'

After the completion of Part 2 four of the colored blocks, red, yellow, green, and blue, were placed before the observer ranging from his left to right in the order named. He was told to lift these blocks separately as many times as he desired, no horizontal rod being used to restrict the height of each lift, and to place them in the order of their weight, the heaviest one to his left, the next heaviest one place to the right, and so on. By means of a stop-watch was recorded the time that it took the observer to arrange these four blocks in the manner described.

All nine of the colored blocks were then placed in chance order before the subject. He was told to select that block which was most pleasing, *i. e.*, whose color was most pleasing, also to select the most pleasing block of the remaining eight, the most and the next most displeasing blocks.

In no case, not even at the close of the test, was the subject told the purpose of the experiment or the point under investigation. No information was given the subject about the results obtained in the experiment.

[1] This varies slightly from the method recommended by Whipple, 'Manual of Mental and Physical Tests,' Baltimore, 1915, Pt. 2, 226, for a similar effect in the case of the size-weight illusion. In his case the graduation while not distinctly mentioned is open to easy discovery. In Part 2 of the present experiment, following upon Part 1, in which no apparent (to the observer) order is followed, no impression is conveyed that the grays are given in any order other than a promiscuous one. Doubtless the repeated giving of the graded series with the colored blocks led to an idea, on the part of the observer, of the gradation, yet in no case, to the knowledge of the writer, was any indication given that the gradation was noticed. With the results of this part of the experiment at hand it seems probable that more reliable results would have been obtained had the subject been informed that the series was a graded one.

The subject's name, sex, year in the university, and the condition of the light were indicated upon the record sheet.

C. Data and Results.—The results obtained by the method of paired comparisons (Part I) are graphically shown in Fig. I.

Arranging the colors in the order of their apparent heaviness we have red, white, orange, violet, green, purple, blue, yellow, black. In this decreasing series three characteristic step effects may be clearly seen. The highest is formed by the two colors red and white; the next step, whose height corresponds fairly closely to the line of chance, may be thought of as made up of orange, violet, green, and purple; the third step is made up of blue, yellow, and black. Stated in another way, the results seem to indicate that red and white have a positive; blue, yellow, and black a negative; orange, violet, green, and purple an indifferent; effect upon the apparent weight of the objects lifted.

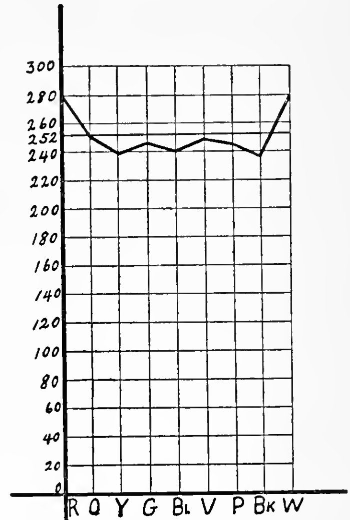


FIG. 1

FIG. I. Method of paired comparisons. The results for 63 subjects are shown. The ordinates represent the total number of judgments of 'heavier' given the different colored blocks.[1] The different colors are represented along the line of abscissas. The height to which a color should rise, assuming the absence of all factors save that of chance, is shown by the horizontal line at ordinate 252.

The average value,[2] the average value in terms of per cent.,

[1] Though no encouragement was given to the subject to reply that the blocks compared were of the same weight such judgments occasionally occurred. In cases of this kind, also in those cases where the subject seemed quite doubtful, the comparison was repeated, but in the curve here shown these cases were disposed of by giving to each color involved one half of one unit credit, it being the plan to give to the color judged 'heavier' one unit credit.

[2] The 'average value' for each color is obtained by dividing the total number of times for all subjects it was judged 'heavier' by the number of subjects. If any color had always received a judgment of 'heavier' regardless of the color with which it was compared its average value would be 8. Assuming no influence other than chance the average value for any color would be 4. The M. V. as given in the table is computed from the average.

and the mean variation for each color are given in the following table, Table I.:

TABLE I

Color	Av.	%	M. V.
Red	4.452	55.6	1.083
Orange	4.008	50.0	0.899
Yellow	3.794	47.4	1.203
Green	3.905	48.8	1.011
Blue	3.810	47.6	1.180
Violet	3.960	49.5	0.943
Purple	3.897	48.7	1.194
Black	3.754	46.9	1.218
White	4.421	55.3	1.194

Table II. is presented to show the gross distribution about the chance line of the individual results obtained. The results, for the men, for the women, and for the men and women are shown in separate columns.

TABLE II

Color	Men			Women			Men and Women		
	$>C$	$<C$	$=C$	$>C$	$<C$	$=C$	$>C$	$<C$	$=C$
Red	12	4	3	22	13	9	34	17	12
Orange	5	9	5	17	15	12	22	24	17
Yellow	4	10	5	15	22	7	19	32	22
Green	8	6	5	13	22	9	21	28	14
Blue	7	8	4	12	20	12	19	28	16
Violet	7	9	3	16	11	17	23	20	20
Purple	8	7	4	14	21	9	22	28	13
Black	5	8	6	14	25	5	19	33	11
White	10	5	4	23	16	5	33	21	9

In the columns headed $>$ C is given the number of individuals whose credit value for the particular color was greater than chance would have allowed. In column $<$ C ($=$ C) is given the number of individuals whose credit value for the different colors was less than (equal to) that allowed by chance.

Considering the totals we note the same step-effect as was pointed out above, and each step is made up of exactly the same colors as before: red and white make up the first; orange, green, violet, and purple, the second; yellow, blue, and black, the third. Even though the number of men tested is quite too small to merit serious consideration of

their results apart from the combined results of both men and women it will be seen that save for orange and blue their distribution is quite in accordance with the statements made above.

If we now note the results obtained from the auxiliary test[1] with the four colors, red, yellow, green, and blue, we have the following table, Table III.:

TABLE III

	Red		Yellow		Green		Blue	
First place	25	13	11	13	10	13	6	13
Second place	16	13	11	13	12	13	13	13
Third place	7	13	17	13	19	13	9	13
Fourth place	4	13	13	13	11	13	24	13

First place (second place, etc.) is used to denote the fact that a block falling therein was designated by the observer as heaviest (next heaviest, etc.) of the four. The first column under each of the four colors indicates the number of individuals giving to the various colors the first, second, third, and fourth place. The second column gives the distribution according to chance.

If we adopt an arbitrary system of units in which the occurrence of a colored block in the first place is given a value of 4; its occurrence in the second place, a value of 3; in the third place, a value of 2; in the fourth place, a value of 1; and compute upon this basis the aggregate value for each of the four colors, we have: red, 166; yellow, 124; green, 125; blue, 105. The figure of chance is 130. These results, save those for yellow, quite accord with our previous statement (see p. 355). To yellow is attached a slightly larger figure than we should have anticipated on the basis of our previous results. To agree with the previous results it should have attained a figure near that of blue, about 105. This slight discrepancy may be due to the presence in this part of the test of complicating factors arising out of the increased freedom in method allowed the subject.

The average time spent in arranging the four blocks in

[1] The method of this test is given on page 354 *supra*.

the order of their apparent weight was 1 min. 59 secs. This result is obtained by averaging the results of forty-seven subjects. The experimenters began to take the time with the sixteenth subject in the series, and failed to get the time in a single subsequent case.

The test requiring a choice of the two most pleasant and the two most unpleasant of the nine blocks led to the following results, Table IV.:

Table IV

	Pleasant		Unpleasant	
Red	2	4	1	4
Orange	4	8	6	9
Yellow	4	4	3	5
Green	7	14	3	4
Blue	14	11	1	4
Violet	7	2	4	9
Purple	6	5	20	11
Black	0	1	12	2
White	6	1	0	2

In the first column under Pleasant (Unpleasant) appears the number of persons choosing red (orange, yellow, etc.) as the most pleasant (unpleasant) color of the nine. In the second column under Pleasant (Unpleasant) is given the number of persons choosing red (orange, yellow, etc.) as the next most pleasant (unpleasant) color. From the table it readily appears that the most generally pleasant color was blue, with green next; the most generally unpleasant color was purple, with black next. Three of these colors, black being the one excepted, showed no particular influence upon the block so colored. Red and white called out little affective quality, and it was these two colors[1] that seemingly had the most marked effect upon apparent weight. Again, of the three colors, yellow, blue, black, which appeared to influence apparent weight negatively, one (blue) is generally pleasing, one (black) is rather displeasing, the other (yellow) is neither decidedly pleasant nor decidedly unpleasant.

The following table, Table V., shows the results obtained from the second part of the experiment (designated above, Part 2):

[1] *Supra*, p. 355.

Table V

Red	68.27	Violet	70.78
Orange	68.32	Purple	70.07
Yellow	67.46	Black	69.57
Green	68.22	White	70.56
Blue	68.30		

The figures opposite the various colors are obtained in the following manner: for each individual test an interpolated value for each color is computed by averaging the value of the last judgment of 'lighter' and that of the first judgment of 'heavier'; the sum of these individual interpolated values is then averaged, and it is the resulting figure that appears in the table. Probably little of a definite nature is to be gained from a consideration of this table. It seems that the method was too crude for satisfactorily disengaging the factor under consideration from the many others which complicate the problem of the apparent weight of a lifted object. Doubtless one of the main complicating factors was the fact that the colored blocks towards the end of the series would not be approached with the same naïveté as were those of the first part of the series. It will be noted that there is a decided tendency to overestimate the weight of the gray block. This overestimation is due, I take it, to the serial arrangement of the grays.[1] It is the result, often noted in the case of unsophisticated observers, and clinging even to the trained observer, of the premature equalization, when there is a problem of equalization, of members of a series of diverse stimuli with a standard stimulus. Seemingly in all such cases there is a tendency to judge that the goal is reached *before* it is reached rather than to judge that the goal is reached *after* it has been reached. This tendency is quite noticeable in the various experiments connected with the problem of the limen. The following is sufficient evidence that this factor was really an effective one in the present experiment. Though each gray block was always heavier than the preceding one very frequently the observer would

[1] Subsequent experiments (Section II.) clearly show that this overestimation is not due to the color (gray) of the block. It may be that the grayness has some influence, but it is relatively small. Compare results given on p. 363 *infra*.

judge a very light block (say one of 64 grams) to be equal to or heavier than the colored block (weight, 76 grams), then judgments of 'lighter,' 'the same,' and 'heavier' would follow. Only one of the sixty-three individuals tested failed to definitely show this tendency, and it was not an uncommon occurrence for this tendency to appear more than once during any particular test. In the opinion of the writer anticipatory attention has entered as a complicating factor in the averages presented. If so, its influence should be stronger for those colors appearing first in the series, since numerous comparisons would in all probability tend to lessen or altogether eliminate it in the case of those colors coming towards the end of the series. Were this factor alone present it seems that the series of values from red to white should show a gradual increase. It is perfectly clear that in the figures this increasing tendency is present. A separate average for the first, second, and third group of three values gives 68.02, 69.10, and 70.07 respectively. But it is quite probable that anticipatory attention is not the only influence present, and the results of our test by the method of paired comparisons, which seems to be the most adequate method of attacking the problem in hand, would lead us to assume an influence of the different colors. The low value (lowest of the nine) of yellow accords with the results previously obtained.

Section II

This section contains a report of additional experiments carried out with the hope of obtaining more satisfactory results relative to the existence of a color-weight illusion. The work was done at Stanford University during the year 1916–1917. The subjects used, sixty-eight men and thirty-one women, were drawn from various classes in psychology, but mostly from my class in mental hygiene. First-, second-, third-, and fourth-year students were represented.

While the apparatus and method employed were similar to those used in the first part of Section I., a brief description is necessary for a clear understanding of the results to follow.

Apparatus.—Six cubes ($2\frac{1}{2} \times 2\frac{1}{2} \times 2\frac{1}{2}$ in.) were constructed

from ordinary Bristol board; the bottom of each was left open; the remaining five sides were covered over with a colored paper, the color varying with the cube. The six colors represented were red, yellow, blue, black, white, and gray. These colors were the same as those used in Section I. with the exception of the gray, which in the present case was No. 17 of Hering's series. (No. 16, the gray previously used, was not available.) Each cube was made to weigh seventeen grams. Adjustment of weight was accomplished by glueing a cork to the bottom of the upper side of the cube and adding shot or paring away cork. The choice of a small absolute weight for the cubes was dependent upon the consideration that if the color of an object has an influence upon its apparent weight that influence should be relatively greater the lighter the object, since the lighter the object the smaller the appreciable difference in weight.

Method.—The subject was seated at one side of a table, the experimenter at the other. The following experiments were carried out:

1. The four cubes, red, yellow, blue, and white, were placed on a large sheet of dark cardboard lying upon the table. The right-to-left arrangement of the cubes varied from subject to subject. The following four orders of presentation were used, beginning at the subject's right: (1) red, yellow, blue, white, (2) yellow, blue, white, red, (3) blue, white, red, yellow, and (4) white, red, yellow, blue. The first (1) represents the order in which the cubes were placed before the first subject; for the next subject the order was changed to that shown in (2); etc.

The subject was told that the cubes were all different in weight and that he should arrange them in the order of their weight, placing the heaviest at his left. He was definitely instructed to look at the weights while lifting them. An occasional prompting was sometimes found necessary to make the subject conform to this requirement. The subject was allowed to rearrange the order[1] as often as desired until

[1] A change in the order often occurred. One of the most frequent means of bringing about such a change was the following: the subject would lift the various

the proper order was obtained. The subject's final arrangement was noted and recorded.[1]

2. A test similar to the one just described was given using the three cubes, white, gray, and black.

3. The subject was tested with cubes red, yellow, blue, white, and black, using the method of paired comparisons. This method has already been referred to under Section I. (p. 351). The same general precautions were taken here as there.

4. In some of the later tests the subject was asked to designate from 'looks' alone the order of the apparent heaviness of the cubes. After the subject had given his order he was asked to explain why the cube so designated looked heaviest and another seemed lightest.

The results obtained for Test 1 appear in Tables VI. (*A*) and VI. (*B*).

TABLE VI. (*A*)

Color	Men (68)				Women (31)				Men and Women (99)			
	1	2	3	4	1	2	3	4	1	2	3	4
Red	18	16	19	15	6	5	13	7	24	21	32	22
Yellow	15	19	22	12	8	10	8	5	23	29	30	17
Blue	17	20	17	14	9	10	5	7	26	30	22	21
White	15	13	10	30	8	6	5	12	23	19	15	42

Under 1, 2, 3, and 4 appears the number of times the different colors were ranked as heaviest, second heaviest, third heaviest, and lightest. The numbers in parentheses following Men, Women, and Men and Women indicate the number of subjects.

Assigning the following values to the different judgments, heaviest, 4; second heaviest, 3; third heaviest, 2; lightest, 1; and computing the aggregate value for each color we have Table VI. (*B*).

The significance of these results is not very clear. They

cubes one or more times and place them in a certain order, then he would compare the extremes, seemingly as a confirmatory test, and that would frequently lead to a change in the established order.

[1] This method will be referred to later as the 'arrangement' method.

TABLE VI. (*B*)

Color	Men (68)	Women (31)	Men and Women (99)
Red	173	72	245
Yellow	173	83	256
Blue	176	83	259
White	149	72	221

do not support the interpretation given to the results of Section I., neither do they accord with the results obtained by the method of paired comparisons referred to later in this section (see p. 364). There seems to be a tendency to judge white and red as relatively light, blue is heaviest, while yellow follows as next heaviest. It is to be noted that there is not complete agreement with the rule: light colors—light weight, dark colors—heavy weight.

The results for Test 2 are given in Tables VII. (*A*) and VII. (*B*).

TABLE VII. (*A*)

Color	Men (68)			Women (31)			Men and Women (99)		
	1	2	3	1	2	3	1	2	3
White	23	16	29	10	5	16	33	21	45
Gray	23	30	15	9	12	20	32	42	35
Black	23	22	23	12	14	15	35	36	38

Under 1, 2, 3 appears the number of times the different colors (grays) were ranked as heaviest, next heaviest, and lightest.

Assigning the following values to the judgments, heaviest, 3; next heaviest, 2; lightest, 1; and computing the aggregate value for each color we have Table VII. (*B*).

TABLE VII. (*B*)

Color	Men (68)	Women (31)	Men and Women (99)
White	186	56	242
Gray	205	69	274
Black	215	71	286

From these results for three blocks differing only in tint it is quite apparent that the tendency is to judge the darker block as heavier. This is true for both men and women.

The results for Test 3 appear in Table VIII.

TABLE VIII

Color	Men (42)	Women (8)	Men and Women (50)
Red	97	23	120
Yellow	90	21	111
Blue	96.5	29	125.5
Black	96.5	30	126.5
White	101	23.5	124.5
Gray	101	25	126

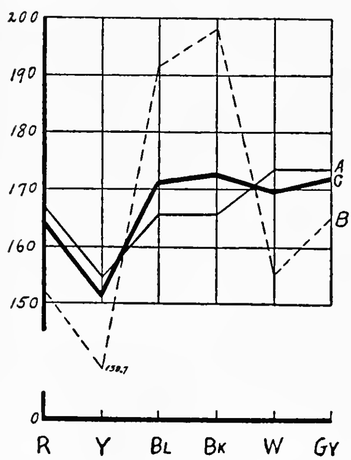


FIG. 2. Curve *A* indicates the relative difference in the apparent weight of six different colored cubes of the same weight, as judged by 42 men. Curve *B*, the same for 8 women. Curve *C* is obtained by combining the results for both men and women. The colors are represented along the line of abscissas, the relative number of judgments of 'heavier' along the line of ordinates.[1] Method of paired comparisons.

The numbers represent the relative number of judgments of 'heavier' given in a comparison of different colored cubes. A graphic representation of these results is shown in Fig. 2.

For the men white, gray, and red are the heavier colors, yellow is the lightest; for the women[2] black, blue, and gray are the heavier, as with the men yellow is the lightest. White is heaviest for the men while for the women it is slightly heavier than black and red.

Since the method of paired comparisons was used in the experiments of both Sections I. and II. it was deemed advisable, for the sake of increasing the number of comparable results from like tests, to combine the values received in the two. The colors common in the two groups were red, yellow, blue, white, and black. The judgments resulting from a

[1] In this figure, as well as in all subsequent ones, the absolute ordinate values appertaining to any individual curve have been transformed so that their summation equals 1,000.

[2] The number of women tested is too small to make sure that their results are typical.

comparison of all possible combinations of these five colors were separated out from the two wider groups of results belonging to Sections I. and II. Table IX. presents the results obtained.

TABLE IX

Color	Section I			Section II			Sections I and II		
	Men	Women	Men and W.	Men	Women	Men and W.	Men	Women	Men and W.
	(19)	(44)	(63)	(34)	(8)	(42)	(53)	(52)	(105)
Red	44.5	109	153.5	69	16	85	113.5	125	238.5
Yellow	34	82	116	67	15	82	101	97	198
Blue	32	78.5	110.5	65.5	17	82.5	97.5	95.5	193
Black	39.5	88.5	128	70	18.5	88.5	109.5	107	216.5
White	40	82	122	68.5	13.5	82	108.5	95.5	204

The table brings together the results, method of paired comparisons, of Section I., Section II., and Sections I. and II. for red, yellow, blue, black, and white. In the different columns appear the values given by men, women, and men

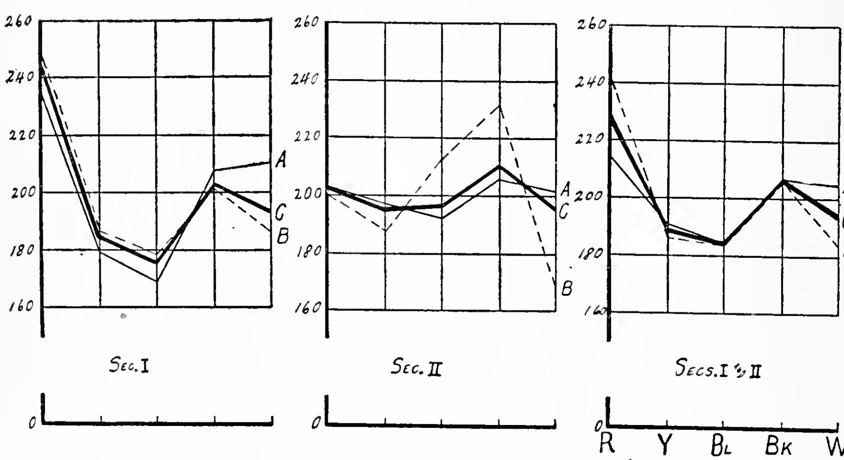


FIG. 3. Method of paired comparisons. The results of Sections I., II., I. and II. combined, are shown graphically. In each case Curve *A* represents the apparent relative weight of the five colored cubes, red, yellow, blue, black, and white as judged by the men. Curve *B*, the same for the women. Curve *C* represents the combined results for men and women. The colors are shown along the line of abscissas; the relative number of judgments of 'heavier' along the line of ordinates.

and women for each color. The numbers in parentheses designate the number of subjects.

For the purpose of facilitating comparison the results of Table IX. are graphically represented in Fig. 3.

It will be seen that there is great similarity in the general direction of the various curves. The results of both Sections I. and II. seem to indicate that there is a general tendency to judge red and black as relatively heavy, white fairly heavy, and yellow and blue light. The main departure from this tendency occurs in the women's judgment of white. For them white occupies a low position similar to that taken by yellow and blue, red and black being the heavy colors. Attention may be called to the fact that for the women red is noticeably the heaviest of all.

The results for Test 4 appear in Table X.

TABLE X

Color	Men (12)	Women (4)	Men and Women (16)
Red	44	15	59
Yellow	23	9	32
Blue	50	16	66
Black	56	23	79
White	25	8	33
Gray	44	13	57

The numbers represent the relative influence of 'looks,' without actual lifting, upon the different colored cubes. These numbers are obtained by assigning the following values to the different judgments, the block judged heaviest, 6; next heaviest, 5; . . .; lightest, 1; and computing the total credit due each color.

A graphic representation of the results appearing in Table X. is shown in Fig. 4.

The results for Test 4 are much more uniform than those received in the other tests. The uniformity from subject to subject rendered the testing of a larger number of subjects unnecessary. The relative position for each of the six colors used is the same for both men and women, save in the case of yellow and white which show a small difference. These results favor the rule: light-colored objects appear light in weight; dark-colored objects appear heavy in weight.

The introspections taken in connection with the judg-

ments of Test 4 indicated quite clearly that the judgments made were due in most cases to associations of a general and not of a specific nature. For example, the connection between the black cube and heavy weight was immediate, because dark objects are usually thought of as heavy, and did not depend upon the arousal of some specific association, like the idea of coal or iron.

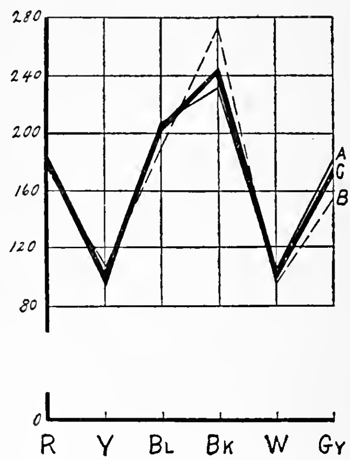


Fig. 4. Curve *A* indicates the relative difference in the apparent weight, judged by 'looks' alone, of six different-colored cubes as judged by 12 men. Curve *B*, the same for 4 women. Curve *C* represents the combined results of both men and women. The colors are shown along the line of abscissas. The ordinates represent the relative weight of the different colors.

We raised an objection to accepting at their face value the results of Part 2 of Section I. because of the probable influence of anticipatory attention. The values for the colors, red, yellow, blue, black, and white (the colors for which we have extended results with the method of paired comparisons), do not vary widely in their distribution from the distribution obtained by the method of paired comparisons. If the value for red were slightly greater (and if we allow for anticipatory attention it would be relatively larger) the curve of distribution for these values would be quite similar to those obtained by the method of paired comparisons. The low position of yellow is to be noted. If anticipatory attention is not admitted the results for blue in the two cases do not accord.

IV. Discussion

The results accruing from the problem herein attacked probably need little further discussion. They indicate that when the method of paired comparisons is used there is a general tendency to judge the apparent weight of red and black as heavier than that of yellow and blue. The fact

that red and black always lie above,[1] yellow and blue below, the line of chance (at ordinate 200) may be interpreted as due to the overestimation of the weight (judged after lifting) of the colors red and black and the underestimation of yellow and blue.

The results received when the arrangement method was used (p. 362 f.) are not in harmony[2] with the above interpretation. They seem to indicate quite clearly that white is judged lighter than red, yellow, or blue.

There is no doubt in the writer's mind that the results obtained by the method of paired comparisons must be considered as more reliable than those obtained by the arrangement method. The former appears to be the more scientific method. A number of distinct judgments is involved and the situation immediately preceding each judgment is comparatively simple. When the arrangement method is used the placing of all the blocks before the subject renders the situation preceding his judgments rather complex. That the method of arrangement permits of rearrangements is probably an argument against it. Often the cause of the rearrangement seems to be extrinsic and not dependent upon the appearance of the cubes (see note, p. 361 f.).

The fact that the results of the two methods, that of paired comparisons and that of arrangement, the cubes being actually lifted, show no definite correlation[3] with the results received when the judgments were made by 'looks' alone seems to indicate that the lifting of the cubes may have introduced an additional factor. One might be disposed to draw the inference that this factor is not entirely dependent upon the appearance of the cubes.

If the claim that the results obtained by the method of paired comparisons demands prior consideration is admitted, the question arises, To what is the differential influence indicated due? The fact that the effect produced is quite analogous to that present in the size-weight illusion suggests the

[1] In Curve *B* of Fig. 3 (Section II.) the ordinate for red is exactly 200.

[2] That is, if they are taken as at all indicative.

[3] With the exception of the results for blue there is a fair degree of correspondence between the curves of Figs. 3 and 4.

possibility of a similar explanation. The size-weight illusion is usually explained on the basis of a close association between large objects and heavy weight, and between small objects and light weight. Making use of a similar argument we should expect yellow and white, since they *appear* light, to be overestimated after lifting. It may be remarked that in the case of white this occurred in the experiments of Section I. But the results for yellow, the most consistent throughout, indicate just the opposite. Further, we should expect the weight of black, blue, and red to be underestimated after lifting, but this is opposite to the general tendency as shown by the results (see curves in Fig. 3). The results, then, militate against an associative explanation so simply conceived.

Since white objects are known, other things being equal, to appear larger than black objects it might be thought that a slight difference in apparent size would exist with our blocks and consequently influence their apparent weight. If the white (black) block is judged larger (smaller) than the black (white) block, and no other factor is involved, it in consequence ought to be judged lighter (heavier) than the black (white) block. This is in accordance with the general tendency of the results for white and black, but does not hold for blue. Even with white the results of Section I. indicate that it was often judged relatively heavy. Again, it has been found that 'there is a small, but decided influence of color'[1] upon the apparent size of an object. According to Quantz[2] red, orange, yellow, and purple increase, blue and violet decrease, the apparent size of an object, green is near the borderline of influence. According to these results the red object should appear larger than the blue or violet object and consequently be judged, when lifted, lighter than they, but this was not found to be the case. A similar argument with respect to some of the other colors shows the inadequacy of apparent size as a basic explanation of the results obtained. Hence we conclude that some factor other than the apparent size of the blocks is involved.

[1] J. O. Quantz, *Amer. J. of Psychol.*, 1895, 7, 40.
[2] *Loc. cit.*, 39.

There is a possibility that the different colors exert different dynamic influence upon the observer such that red and black, say, call out but little neuro-muscular force, and blue and yellow call out more of such force. This idea is quite too conceptual to be introduced into the present discussion, which already has involved too much theorizing.

V. Conclusions

The writer is fully aware that the conclusions to be drawn from the results herein presented must necessarily be tentative. It is perfectly clear that for a final solution of the present problem much careful and accurate work must be done. At the present time I only wish to suggest the following conclusions as sufficiently justified by the results obtained:

1. The influence of the color of an object upon its apparent weight is relatively slight.

2. There is a tendency in many cases, at least when the method of paired comparisons is used, to judge a red or black object to be slightly heavier than a yellow or blue object of the same weight.

3. Apparently the influence of these colors is not due to their tint value alone; it seems necessary to consider their hue as a minor factor.

4. The results obtained do not lend themselves to a simple associative explanation.

5. Seemingly there is no simple correlation between the affective quality of a color and its influence upon apparent weight.

The writer wishes to express his indebtedness to Professor Warner Brown for a helpful criticism of an early draft of this paper.

TACTUAL ILLUSIONS OF MOVEMENT[1]

BY HAROLD E. BURTT

A. Introduction

The kinematoscopic illusion in which two similar stationary visual stimuli in quick succession in different places yield an impression of movement has long been familiar. Wertheimer[2] and Korte[3] studied the phenomenon rather exhaustively with variations of intensity, distance, interval, etc. The writer has demonstrated the possibility of such illusions in audition.[4] Benussi has found similar phenomena in touch and has published two brief accounts,[5,6] the latter of which appeared after the present study was in progress. It seemed profitable to determine whether the tactual corresponded to the other sensory fields merely in the existence of the illusion under optimal conditions, or whether there was a correspondence in the more detailed aspects. Accordingly an investigation was made of the rôle of the intensity of the tactual stimuli, the distance between them and other variables much as they were studied by Korte in the visual phenomenon.

The principal facts found by Korte are as follows:

1. If two equal straight lines or other similar visual stimuli a few centimeters apart are presented in quick succession for equal lengths of time certain conditions of intensity, time and distance yield an impression of movement from one

[1] From the Harvard Psychological Laboratory.

[2] Wertheimer, M., 'Experimentelle Studien über das Sehen von Bewegung,' *Zeitschrift für Psychologie*, 1912, 61, pp. 161–265.

[3] Korte, A., 'Kinematoskopische Untersuchungen,' *Zeitschrift für Psychologie*, 1915, 72, pp. 193–296.

[4] Burtt, H. E., 'Auditory Illusions of Movement,' J. of Exp. Psychol., 1917, 2, pp. 63–75.

[5] Benussi, V., 'Kinematohaptische Erscheinungen,' *Archiv für die gesamte Psychologie*, 1913, 29, pp. 385–388.

[6] Benussi, V., 'Kinematohaptische Scheinbewegung und Auffassungsumformung,' *Bericht VI. Kongres für experimentelle Psychologie*, Göttingen, 1914, pp. 31–35.

stimulus to the other in the direction of the actual temporal succession.

2. The longer the exposure of the stimuli the relatively shorter the interval which gives the optimal impression of movement.

3. The greater the distance between the stimuli the greater the optimal interval.

4. The greater the absolute intensity of the stimuli the shorter the optimal interval.

5. The greater the intensity the greater the optimal distance.

6. If the second stimulus is of greater intensity than the first the direction of the apparent movement is sometimes the reverse of the actual direction of temporal succession.

The above relations were investigated in the present case with tactual stimuli. The effect of varying degrees of difference in intensity between the two stimuli, which occupies a considerable portion of Korte's monograph, was not studied inasmuch as the reversal effect occurred much less frequently in the present case.

B. Apparatus and Method

An experiment of this sort necessitated some method of giving two similar tactual stimuli at different points of the skin, with control of the time relations, intensity of the stimuli and distance between them. It was also desirable to give as a check some stimulus approximating in its effect a continuous movement on the skin.

Ten brass rods with blunted points 2 cm. apart in a straight line were arranged to stimulate the back of the forearm. With the whole row actuated by solenoids in quick succession the effect was always much like a continuous movement, for any two adjacent points were within a "Weber's circle." This gave the check mentioned above. For most of the work only two solenoids were actuated on a given trial, the distance between stimuli depending on which solenoids were used.

The solenoids were mounted individually on narrow strips

of wood arranged to slide vertically side by side on a large wooden base. A piece of round magnet iron 2.5 cm. long formed the core and was suspended by a spiral spring from an adjusting screw above the solenoid. The lower end of this core contained a hole into which a brass rod of the same diameter as the core was dowelled. The lower end of this brass rod was pointed to give the tactual stimuli. The subject's arm was clamped palm downward in the rest of the Mosso's ergograph with the first and third fingers in the usual tubes to insure rigidity. The base to which were attached the strips of wood carrying the solenoids was mounted longitudinally above the arm. These strips could be moved up and down individually by set screws. The base also could be moved vertically so it was possible to adjust the 10 brass points to any desired distance from the skin. Any hairs in the vicinity of the brass points were of course removed. The iron part of the cores were usually set a few millimeters above center and the brass points adjusted to a distance of about 1 mm. from the skin. Thus the intensity of the stimulus did not vary appreciably with difference in length of time the current passed through the coils although it did vary with the intensity of the current. This latter was controlled by rheostats. The solenoids were actuated by either about 20 volts of storage batteries or by the direct 110-volt laboratory current with a rheostat in series.

A time-controlling mechanism was made from the 'Leipzig time sense apparatus' with a series of closely adjacent marginal contacts and a rather complicated switchboard. It was possible to actuate the 10 solenoids in rapid succession or any two in succession for desired lengths of time with any desired interval between. The space order (up or down the arm) could be reversed at will. The time mechanism was placed in an inner soundproof room. The experimenter and subject sat in an outer room at opposite sides of the table on which were the solenoids, switchboard and rheostats. The subject's arm rested comfortably in the apparatus which was so placed that the upright base was between the subject and the solenoids. Their action was thus invisible

to the subject but could be watched by the experimenter. The action of the solenoids was practically noiseless.

The subjects were not told of the nature of the experiment but merely that it was a study of tactual perception. At the beginning of each hour they were instructed: "When I say ready close your eyes and attend to the forearm and then tell what you felt."

The warning 'ready' was given approximately 1 second before the stimuli. On a given hour trials involving different relations of the given variables were presented in irregular orders. The main interest was in the effect of two successive stimuli, but trials were always intermixed in which a row of adjacent stimuli were given. It could thus be seen how the subject would confuse two discrete stimuli with the close approximation to continuous movement produced by a row.

The experiments were performed in the Harvard Psychological Laboratory in the academic year 1915–16. The subjects were graduate students or undergraduates of considerable psychological experience.

C. Results

I. *The Movement Illusion*

All four subjects who participated in the experiment yielded in many trials the illusion of movement. If the point of a rod was pressed on the skin of the forearm for a definite period and then after a brief interval another rod a few cm. distant pressed on the skin for the same length of time as the first, there was sometimes an impression of movement from the first point to the second, sometimes a single fused sensation and sometimes two discrete impressions. These effects varied with the combination of time, intensity and distance variables employed (*cf.* infra).

The (illusory) movement impression was variously characterized in the introspection of the subjects.

Subject *A* distinguished a *series* in which there appeared to be a number of stimuli in a row 'pretty close but possibly felt as separate,' from a continuous *flow* from one end to the other in which he 'could not tell the separate points.'

There was also at times an *arrow* at the end of the flow as if the movement went on in the air a short distance.

C noted three forms of the illusion: a *line* which was 'as if you pressed down a long object' either all at once or beginning at one end of the object; a *walk* in which there are a number of discrete points in succession; and a *roll* which is 'continuous movement.' He also occasionally reported a *loop*, the motion being in the air above the arm rather than actually upon the skin.

D described the apparent movement as a *snake effect.* He felt a number of intermediate points but also a connection. It suggested black points connected by a line; 'when they were close together the line was thick and when far apart it was thin.' The effect was best when all points appeared equidistant. A frequent report was three points in succession in the order *ABC* or *ACB.* In the latter case, with the middle point felt last, the snake effect appeared less readily.

F reported movement in which they 'skip along like playing four or five notes on a piano in succession.' It often suggested a little boy running along. He also noted frequently a sort of *loop* in the air between the ends. This 'blends into a sensation of movement.' This loop reported by *C* and *F* corresponds to Benussi's 'Bogenbewegung in der Luft.'[1] Benussi found that with a greater time interval the Bogen was greater and the movement worse. This seemed true for the present subject *F.*

II. *Exposure and Interval*

Korte found in his visual studies that the longer the exposure of the stimuli the relatively shorter the interval that yielded the optimal impression of movement. The writer found the same to be true with auditory stimuli.

To test this factor in the present case, intervals between stimuli of 15, 21, and 40 sigma were used with exposures of the same length or multiples thereof. The distance between the stimuli was constant at 12 cm.; the intensity was moderate, that given by a current of 1.5 amps. in the solenoids.

[1] *Op. cit.*, p. 32.

The results are summarized in Table I. The three successive columns represent results obtained with intervals of 15, 21 and 40 sigma respectively. At the left of each column are arranged vertically the exposures (in sigma) which were

TABLE I

EXPOSURE AND INTERVAL

(Distance 12 cm. Intensity 1.5 amps.)

Subject	Exposure	Interval 15	Exposure	Interval 21	Exposure	Interval 40
A			126	Two ends more discrete		
			105	Two ends	200	Two ends rather close together
			84	Series or two	160	Short series then other end
	45	Series or flow	63	*Series or flow*	120	Series three or four
	30	Two adjacent	42	Slight flow	80	*Four in flow or series*
	15		21	Single one in middle	40	Three in series
C			147	Two or three at each end		
	90	Ends successive	126	Five rolling with gap	240	*Four to six in succession*
	75	Ends successive	105	*Six or seven in succession (walk)*	200	Three or four in succession
	60	Four or five in succession (line) or two ends	84	One end then three rolling at other end	160	Three in succession
	45	*Four in succession (line)*	63	Three in succession	120	Ends successive but spread
	30	Ends successive	42	Ends successive or simultaneous	80	Ends successive but spread
	15	Ends successive or simultaneous	21	Ends successive or simultaneous	40	Ends successive
D			126	Group at ends		
			105	Two or three in series	200	Several at each end
			84	*Three or four with heavy line*	160	Three or four in series
			63	Three with thin line	120	Three or four in series
			42	Several simultaneous	80	*Four with snake effect*
					40	Three or four in series or ends
F	60	Ends	84	Ends		
	45	Ends	63	*Three or four*		
	30	*Three or four*	42	Two or three		
	15	Two adjacent	21	Two close together		

combined with the given intervals. In each section is stated briefly the phenomenon that generally occurred with that given combination of exposure and interval for the subject

indicated in the margin at the left. Each statement is based on from 5 to 10 trials and all results in the table were obtained on the same day. The check trials in which an actual series of touches was given by several successive solenoids are omitted. The results in this and subsequent tables are for only those trials in which two stimuli were given for equal lengths of time with an interval between. For instance we note that with the stimuli for 30 sigma each with an interval of 15 sigma between them, subject *A* generally reported two adjacent touches, and with stimuli for 42 sigma each and an interval of 21 sigma he noted usually a 'slight flow.'[1] Following across any horizontal row in the three columns we have exposure and interval in the same ratio. Thus in the example above 30 : 15 : : 42 : 21 : : 80 : 40. That report in each column which seems to indicate the greatest effect of the illusion is indicated in italics. Thus by noting the distribution of the italicized phrases for the different subjects any individual or general tendencies may be seen. The variables are arranged (in all tables) with the lowest values at the lower left corner so that an array of italics extending obliquely upward to the right indicates a positive relation between the variables and an array extending obliquely downward to the right indicates an inverse relation.

In the table it is evident that two of the subjects show a direct relation and two an inverse relation between the exposure and interval which give the best impression of movement. For instance with subject *A* an interval of 21 sigma is best combined with exposures of 63 sigma, 3 times as great, while a longer interval (40 sigma) is best combined with one only twice as great (80 sigma). That is, the longer the interval the relatively shorter the optimal exposure. *D* shows similar results. *C* and *F* manifest the opposite tendency but the latter's difference is very slight (a matter of 3 *vs.* 4 apparent stimuli) and perhaps negligible. On the whole the difference appears slightly in favor of the inverse relation. The longer the exposure of the stimuli, the relatively shorter the optimal interval. One may note further

[1] For explanation of individual notations, see p. 374–5.

that for a given interval the shorter exposures give an impression of simultaneity or fusion into a single touch, while increase of the exposure produces the movement illusion and still further increase an impression of discrete successiveness, a result in accord with Korte's visual findings.

III. *Distance and Interval*

Table II. summarizes the results with the distance between the points on the skin and the time interval between the stimuli constituting the variables. The exposures were constant at 60 sigma and the intensity at 1.5 amps. through the solenoids. The successive columns give results with the lengths of interval indicated. The rows give the different distances employed. The general tendency for each combination of interval and distance is indicated by a phrase as in Table I., and the report in each row showing the clearest evidence of the illusion is italicized, where such evidence is present. The results for *F* are indicated numerically rather than qualitatively. He usually gave the numbers of the points which he thought touched the skin (having adopted a scheme with No. 1 next the wrist, etc.). From three to six points were reported in each trial. These were simply added and averaged and the figure in the table represents the average number of successive stimuli reported per trial on the given combination when only two stimuli were actually given.

A glance at the italics shows in every instance a tendency toward a distribution obliquely upward to the right. That is, to obtain the optimal impression of movement the longer the interval between the stimuli the greater must be the distance. This result is the same as that found by Korte in vision. We may note further that for a given distance small intervals give a single point or simultaneity, while increase of interval gives movement and later discreteness and succession.

IV. *Absolute Intensity and Interval*

The intensity of the stimuli was varied from 1.8 amps. through the solenoids, which gave a moderate touch to 4.4

Table II

Distance and Interval

(Exposure 60 sigma. Intensity 1.5 amps.)

Subject	Distance	Interval: 15 Sigma	30 Sigma	45 Sigma	60 Sigma	75 Sigma	105 Sigma
A	16 cm.		Ends or flow at each end	Ends	Ends	Ends	Ends
	12 cm.		1 spread out	2 adjacent with arrow	2 or 3 adjacent	*3 in series*	
	8 cm.		1 or short flow	1 or short flow	*Flow*	Flow or series	Series of 3 or 4
A (later)	16 cm.	Ends	Ends	Ends	Ends	Ends	
	12 cm.	Short series or flow	Series or 2 adjacent	*Flow or series*	Two	Two	
	8 cm.	*Flow*	Series	Series or 2 adjacent	2 adjacent	2 adjacent	
C	16 cm.	Ends successive or simultaneous	Ends or series	Ends or series simultaneous	Ends each spread	Ends each spread	Ends
	12 cm.	Ends	Ends	2 or 3 successive or simultaneous	*4 in succession*	*4 in succession*	Ends spread
	8 cm.	Several in succession or simultaneous	3 in succession	*3 or 4 in succession*	*3 or 4 in succession*	3 in succession	3 in succession
D	16 cm.	Ends	Ends	Ends	*Slightly snaky*	Ends	
	12 cm.	Several simultaneous	Several simultaneous	*Snake*	*Snake*	2 or 3 in series	
	8 cm.	One	One	One	One	2 near together	
F	16 cm.	3.2		*4.8*		4.0	
	12 cm.	*7.2*		5.8		4.8	
	8 cm.	*5.0*		4.2		4.4	

(Exposure 120 sigma. Intensity 1.5 amps.)

		40 Sigma	80 Sigma	120 Sigma	160 Sigma
C	16 cm.	Discrete groups	*Discrete groups walking or rolling*	Ends slightly spread	Ends slightly spread
	12 cm.	*5 or 6 rolling*	4 or 5 walking or line	Discrete groups	Discrete groups walking
	8 cm.	Several walk	Several walk or roll	Several walk with gap	Several walk

amps., a stimulus so intense as to be startling. In the separate columns of Table III. are the results with different intensities. The rows give the intervals. The exposure was held constant at 60 sigma and the distance at 12 cm. The

Table III

Intensity and Interval
(Distance 12 cm. Exposures 60 sigma)

Subject	Interval	1.8 Amps.	2.1 Amps.	2.4 Amps.	3.2 Amps.	4.4 Amps.
A	75 sigma	Two adjacent	Two adjacent	Ends	Ends	Ends
	45 sigma	*Two adjacent or flow or arrow*	*Two adjacent or flow or series*	Two or series	Two quite close	Two farther apart
	15 sigma		Series or flow	*Flow*	Two adjacent	Two farther apart
C	75 sigma		Several in row	*Several in row or line*	Ends or row with loop	Ends or row with loop
	45 sigma		One or two	*Several in row sometimes loop*	Several in row	Several in row or ends
	15 sigma		Several in row or one	Several in row	*Series with loop or lot of motion*	Series or ends
D	75 sigma		2.0	2.2	3.4	*3.8*
	45 sigma		1.6	2.4	3.0	*3.4*
	15 sigma		1.6	2.0	2.6	*3.8*
F	75 sigma		4.2	*4.8*	3.8	4.4
	45 sigma		4.4	4.6	*5.2*	*5.2*
	15 sigma		4.4	6.6	5.6	*7.0*

general tendency shown on each combination is briefly indicated as in previous tables and the phrase in each row which indicates the optimal illusion italicized. The results for *D* and *F* are recorded numerically as in Table II.—the average number of touches felt per trial.

The italics are distributed with three of the four subjects in a direction obliquely downward toward the right. With *D* the greatest movement occurred always with the greatest intensity. Thus the general tendency of the majority is for a greater intensity to require a lesser interval to produce the movement impression. Further, for the two subjects with qualitative results, with a given interval as the in-

tensity increases the impression passes from that of a single touch, through a row or motion to two discrete touches. Both these facts agree with those found by Korte.

V. *Distance and Intensity*

In Table IV. the distance between the stimuli and the intensity of the stimuli were the variables. The times of exposure and interval were constant, in most cases 60 and 30 sigma respectively. In a few series other times were used, as indicated. The successive columns indicate intensity of the stimuli in terms of the current in the solenoids. The rows for any subject indicate the distance between the stimuli. The most frequent report for each combination of distance and intensity is indicated as previously. In a given row the report which shows the most marked movement illusion is italicized. *F*'s results are given in the form of the previous tables, *i. e.*, the average number of touches felt per trial.

The array of italics for the various subjects shows a rather general tendency toward a distribution extending obliquely upward toward the right, *i. e.*, with a greater distance between stimuli a greater intensity affords a better illusion of movement. This too is in accord with the results of visual experiments. Further, for a given distance, as the intensity varies from weakest to strongest there is somewhat of a tendency for the effect to pass from unity through movement to discreteness, and for a given intensity increase of distance tends to produce increase of discreteness, a result similar to Korte's. In this latter connection we may note Benussi's finding that with distances greater than 16 cm. on the forearm there was an individualizing of the ends following the superficial movement.

VI. *Unequal Intensity*

In all the experiments up to this point the two successive stimuli on a given trial were equal in intensity. This factor was now varied as had been done in the visual and auditory experiments. With exposures of 60 sigma and interval of

Table IV

Distance and Intensity

(Exposures 60 sigma. Interval 30 sigma)

Subject	Distance	1.2 Amps.	1.4 Amps.	1.5 Amps.
A	12 cm.	Ends	*Flow or series*	*Flow or series*
	8 cm.	*Flow or series*	Series	Ends or series
C	12 cm.	One or both ends	*One end rolling*	Ends sometimes spread
	8 cm.	*Three rolling*	Three rolling or walk	Ends sometimes spread
D	12 cm.	Ends	*Ends or several simultaneous*	Ends
	8 cm.	One or snake	*Good snake*	Two or snake
F	12 cm.	4.4	5.0	3.8
	8 cm.	4.2	4.4	4.4

Subject	Distance	2.1 Amps.	2.3 Amps.	2.8 Amps.	3.6 Amps.	4.4 Amps.
A	16 cm.	Ends close		Ends		Ends
	12 cm.	Ends or flow or series		*Flow or series*		Ends
	8 cm.	*Ends or flow*		Ends		Ends
D	16 cm.	Ends or one		Ends		*Ends or series*
	12 cm.	One or series		*Series*		*Series*
	8 cm.	One or two		*Long series*		Series
C	16 cm.	Ends	Ends	Ends	*Ends with connection*	
	12 cm.	Two adjacent	*Two close, with loop*	Two at one end	Two adjacent	
	8 cm.	One or two	*3 or 4 successive*	One or two	One or two	

(Exposures 60 sigma. Interval 45 sigma)

Subject	Distance	2.1 Amps.	2.3 Amps.	2.8 Amps.	3.6 Amps.
C	16 cm.	Ends or one	Ends with line on skin	Ends with or without loop	*Ends with loop*
	12 cm.	One or two	One or two	*2 or 3 with loop*	Two connected or with loop
	8 cm.	One	*One or two with loop*	One or two	One

30 sigma, with stimuli 8 or 12 cm. apart and with the intensity of the first stimulus 1.5 amps., that of the second was increased from 0.5 to 2 amps. It was varied irregularly through various intensities and trials were intermixed in which both were equal.

Table V. gives the per cent. of trials in which the apparent movement occurred in the direction the reverse from that of the actual temporal succession. The first column gives the per cent. of trials in which the reversal effect occurred when the second stimulus was more intense (including all variations

TABLE V

UNEQUAL INTENSITIES

(Exposures 60 sigma. Interval 30 sigma. Distance 8 or 12 cm. Intensity 1.5 amps.)

Subject	Per Cent. of Trials Yielding Reverse Movement: Second Stronger	Equal Intensity
A	10.5	0
C	21.4	0
D	32.0	0
F	13.1	0

of intensity); the second column gives the same figures for the trials with equal intensity of the two stimuli. The results were obtained on several days and are based on some 50 trials of each sort for each subject.

It is evident that whereas with equal intensity of the stimuli the reversal effect never occurs, with the second stimulus more intense by various amounts the effect occurs in from 10 to 32 per cent. of the trials with the different subjects, an average of 19 per cent. This per cent. is much less than that found by the writer in the auditory illusions.

No attempt was made to correlate the results with the extent of the intensity difference or to study other aspects of the reversal effect (Korte's delta movement) as was done in the visual studies for the simple reason that the reversal effect occurred so much less frequently. However it is significant that the effect does exist there just as it does in vision and audition.

D. CONCLUSIONS

The above experiments investigated in the tactual illusion of movement the most salient of the factors noted by Korte in the similar visual illusion. The following are the principal results:

1. Two punctate tactual stimuli on the forearm for equal lengths of time separated by a discrete time interval and a

few centimeters apart, yield under certain conditions of time, distance and intensity an impression of movement from one point to the other in the direction of the actual temporal succession.

2. To produce the optimal impression of movement the following relations are required between the different variables under consideration:

(*a*) The relation between the length of exposure of the two stimuli and the time interval between them is not especially clear cut but on the whole is inverse—the longer the exposure the relatively shorter the interval.

(*b*) The greater the distance between the stimuli the greater the time interval.

(*c*) The greater the intensity of the stimuli the less the interval.

(*d*) The greater the distance between the stimuli the greater the intensity.

3. If the intensity of the second stimulus is greater than that of the first the illusory movement is sometimes produced in the reverse direction.

The above results are quite similar to those found by Korte in the visual illusion of movement. They are also the same as those found by the writer in a preliminary study of auditory illusions as far as that investigation was carried.

It seems a significant fact that the same illusion of movement is found in vision, audition and touch and that the same laws as to the relation between the variables involved hold to a considerable degree in the three fields. This points of course to a central origin. Whether it is due as Wertheimer claims to a physiological 'Kurzschluss' between cortical areas corresponding to the two stimuli at different points in space is an open question. Such a correlation between points of external space and points of the brain is perhaps plausible in vision and touch. The writer has pointed out[1] that such a correlation is highly improbable in audition and that hence a different explanation is necessary, and inasmuch as the three sense departments show the same phenomena it is better to adopt a single theory applicable to all three.

[1] *Op. cit.*, p. 75.

The writer has suggested (*op. cit.*) that the illusion is due to a continuity of motor impulse, and such a theory would be applicable to touch as well as to vision and audition. A tactual stimulus on the forearm produces an impulse to make some motor adjustment to bring that region into the focus of attention, presumably an impulse to turn the head or eyes toward that point. This would have, of course, a biological significance. A second stimulus, farther up the arm for example, gives a second impulse, presumably in the same centers of the motor cortex, to turn the head or eyes still farther. If the interval is of such length that the second impulse supervenes just when the first is about exhausted this continuity of motor impulses gives the movement illusion. When the second follows too quickly the two are not differentiated sufficiently and there is fusion and if the second is too late there is discontinuity and discreteness. Benussi found it impossible to obtain certain of the phenomena when the stimuli were applied to the forehead. This fact is suggestive inasmuch as the usual motor adjustments of the head to bring that region into the focus of attention are impossible.

The various relations between the different variables necessary to produce the optimal illusion are as explicable by such a theory as by the Kurzschluss theory. If we assume that intensity, distance and length of exposure are related positively to the temporal facilitation of the motor impulses and to their consequent sooner subsidence, all the above phenomena can be readily explained. The writer reiterates his suggestion that the action theory can account for the movement illusion which occurs similarly in vision, audition and touch.

ASSOCIATION-REACTION AS A TEST OF LEARNING[1]

BY KNIGHT DUNLAP

Since 'mental' learning involves primarily the formation of new associations, it is obvious that there is the possibility of testing the learning by some form of the *association reaction*. Moreover, since we can control the associative recall by the proper instructions, it would seem possible to determine the amount of intelligent learning (intellectual association) which has occurred, excluding rote-learning, or association by contiguity: and, by careful selection of controls, it might be possible to determine whether the student has acquired the proper habits of thinking in a given subject-matter, as discriminated from a mere acquaintance with details. Although theoretically possible, the practicability of such a learning-test has remained to be decided by experimental work. I have had such experimentation in mind for some time, and have this year, through the assistance of eight members of my 'teachers' course' in introductory psychology, been enabled to make a beginning in the work. Although a mere beginning, the results have been so decidedly encouraging, that it seems advisable to describe the tests for the benefit of any who may be interested in the same line of work.

The value of a test of the sort undertaken depends on the student's attitude towards it, especially if it is desired to compare the results with the results of ordinary written examinations and class-work. Hence, I announced to my class that any members who were willing to let their standings in the second half-year's work depend on the results of an association-test, would be permitted to take this test in lieu of the regular written examination. Nine students volunteered under this condition, which makes the results com-

[1] From the Psychological Laboratory of the Johns Hopkins University.

parable with ordinary ratings of these students, in so far as the seriousness of their attitudes was concerned.

With careful reference to the assigned text, and the lectures, of the semester, the following list of stimulus words was prepared, each word being accompanied by appropriate control instructions. The significance of these words and instructions depends, of course, on the way in which the materials were presented in class, and hence cannot be adequately explained here. The instructions for each stimulus word were read to the reactor just preceding the giving of the word; and because the instructions were not adequate, they were supplemented by additional extemporaneous instructions, sufficient to make clear to the reactor the relations involved in the control; yet without giving him more information than is actually intended by the written instructions. Extemporizing in this way is of course a fault, and inevitably leads to unfairness, however carefully it may be done. With more carefully selected words and more skillfully prepared instructions, supplementation and its evils can be excluded.

It was made clear to the reactors that the *response* under any given control need not satisfy the grammatical relations of the instruction so long as it clearly indicated the matter which, under the instructions, was demanded by the stimulus word. It was explained that in every case a satisfactory response could be given in one word, but, that in case a logically valid response consisting of a phrase were given, it would be accepted. The reactor was instructed that his primary duty was to give the correct response, but that it was to be given as quickly as possible; that if he thought of the correct answer after giving an incorrect one, he should give the correct one also.

The reaction times were taken by means of the Johns Hopkins chronoscope[1] and the voice keys[2] of the new (smaller) model. The chronoscope, running on a 50 vibration fork, read in units of 2σ.

[1] J. of Experimental Psychology, Vol. 2, No. 3.

[2] Psychol. Rev., 1913, **20**, 250–253.

Instructions and Stimulus Words

1. Reply with the species under the genus I name: that is, give one sort of whatever I name. (Examples: Fruit—Peach; Crime—Murder.) *Consciousness.*

2. Reply with another member of the same class to which belongs whatever I name: give something of the same sort. (Examples: English—Italian; Chair—Table.) *Thyroid.*

3. Same instructions. *Ideo-motor.*

4. Reply with the name of the gland which controls the process I name. *Growth.*

5. Reply with the essential factor involved in that which I name: that which makes it what it is. *Memory.*

6. Same instructions. *Will.*

7. Same instructions. *Judgment.*

8. Reply with the emotion which, in regard to the inclusion of certain elements, is the opposite of the one I name. *Religious.*

9. Reply with the affection (affective content) which in its occurrence or behavior is characterized by that which I name. *Regression.*

10. Reply with that which in regard to *time* is the opposite of what I name. (Example: Yesterday—Tomorrow.) *Anticipation.*

11. Reply with that which is essential to that which I name and whose existence is demonstrated by the latter. *Recall.*

12. Reply with the most important factor involved in that which I name. *Emotion.*

13. Reply with one of the stages in that which I name. *Learning.*

14. Reply with one of the two sorts of that which I name. *Ego.*

15. That which I name depends on a peculiarity of consciousness which is exhibited in all perception. Reply with this peculiarity, or the more general manifestation of this peculiarity. *Rhythm.*

16. Reply with the opposite of that which I name (not the verbal opposite, but that which is actually opposite). *Instinct.*

17. The most fundamental psychological problem is in the relation between that which I name and a second fact. Reply with this second fact. *Reflex.*

After comparing the results of the eight reactors, it seemed possible to evaluate them according to the following rules.

1. Each correct response, given in 900 to 1,000 units, was assigned ten points credit.

2. For a correct reply in 800 to 900 units, one point; in 700 to 800, two points; 600 to 700, three points; and 500 to 600, four points additional credits were given.

3. For a delay from 1,000 to 1,200 units, one point; 1,200 to 1,500, two points; 1,500 to 2,000, three points; 2,000 to 3,000, four points; 3,000 to 5,000, five points, were subtracted.

4. Answers not completely satisfactory, but having a real bearing on the requirements were evaluated at less than ten. (In a properly prepared test, no such cases should arise.)

5. The total credits so obtained were divided by the 'normal' of 170.

The responses obtained from the nine subjects are given below, together with the credits assigned for the normal time (900–1,000 units). Where no credit is noted, it was zero. The numbers in parentheses indicate the number of reactors giving the same response: where no such number is given, 1 is understood.

RESPONSES

1. 'thought' ten points; 'present' ten points; 'content' (2); 'sensation;' 'reflex-arc;' 'sub-consciousness;' *no response* (2).

2. 'adrenal' (3) ten points; 'pituitary' (4) ten points; 'thymus' ten points; 'gland.'

3. 'perceptual' (2) ten points; 'volitional' ten points; 'response;' 'ideational;' 'motor;' 'gland;' *no response* (2).

4. 'thyroid' (6) ten points; 'pituitary' (3) ten points.

5. 'recognition' (3) ten points; 'retention' (2); 'perception' (2); 'reproductive imagination;' 'recall.'

6. 'desire' (4) ten points; 'volition—desire' eight points; 'action' six points; 'volitional' (2); 'consciousness.'

7. 'concept' (5) ten points; 'consciousness' (2); 'thought;' *no response.*

8. 'atheist' 'doubtful;' 'non-religious;' 'anti-religious;' 'sacrilegious;' 'secular;' 'faith;' *no response* (*2*).

9. 'conation' (2) ten points; 'conative' ten points; 'repugnance' ten points; 'progression;' 'retrogression' 'association;' *no response* (*2*).

10. 'memory' ten points; 'realization;' 'disappointment;' 'hope;' 'future;' 'looking backwards;' *no response* (*3*).

11. 'retention' (2) ten points; 'association' (2) eight points; 'memory;' 'perception' (2); 'perceive;' 'experience.'

12. 'cognitive elements' ten points; 'bodily sensation' (3) ten points; 'object' (meaning 'the cognitive element') ten points; 'feelings—physiological activities' ten points; 'feeling' ten points; 'hedonic tone' ten points; 'pleasure.'

13. 'imitation' (8) ten points; 'recall—memory.'

14. 'self' (3) ten points; 'conscious;' 'sub-conscious;' 'selfish—egotistical;' *no response* (*3*).

15. 'specious present' (5) ten points; 'periodicity—specious present' ten points; 'time;' 'feeling;' *no response*.

16. 'habit' (3) ten points; 'learning' (2) ten points; 'consciousness;' 'volition;' 'memory;' 'emotion.'

17. 'consciousness' (5) ten points; 'integration' ten points; 'instinct;' 'action;' *no response*.

As may be inferred from the list, the test was somewhat too difficult. The reactors graded from 90 per cent. to 30 per cent., when they should have graded, according to class-work and previous tests, from 100 per cent. to 60 per cent. The 'order of merit' and relative standings in the test were however what might have been expected from previous work of the students, with one exception: one student who made the low mark (30 per cent.) confessed afterwards to have slighted the reading in the second semester, and to have done no reviewing; this was done with full intention, from curiosity to 'see what the test would show.' Two of the students had reviewed with especial regard to the test; attempting to strengthen word-association in the various topics which they conjectured might be of service, basing their conjecture merely on a general acquaintance with the association-reaction method. These conjectures of course gave the students no help at all on the actual test.

The results of the test brought out sharply the difference between topics which had been thoroughly discussed in class, and those (*e. g.*, 9) which had been slighted, and also difference in difficulty in the topics themselves. As regards the individual differences of the reactors, the test alone does not show whether these were due to differences in amount of work done, or to difference in the ability of the student to grasp the vital points, with an average amount of effort. It is probable, however, that a more carefully prepared test would differentiate these causes. One thing which stands out sharply is that 'bluffing' is entirely excluded: the student makes a good or poor showing, according as he 'knows' or does not know the subject, whatever the contributing cause of his grasp or lack of grasp.

As in all association-tests, the most vital point is in the construction of the word-tests, and the instructions. A special difficulty, which must be solved experimentally, is the derivation of a just and satisfactory method of evaluating the results of the test.

Journal of
Experimental Psychology

Vol. II, No. 6. December, 1917

UNIDEXTRALITY AND MIRROR-READING

BY JUNE E. DOWNEY AND EDWIN B. PAYSON

University of Wyoming

The suggestion out of which the present investigation grew was the possible correlation of expertness in mirror-reading and degree of one-handedness. This possibility has been indicated as a tentative conclusion by the outcome of a previous investigation.[1] In many ways such a conclusion is highly stimulating, since it suggests relationship between what might seem to be a purely sensory capacity on the one hand, and a motor function on the other. The capacity of individuals for mirror-reading varies very greatly as shown by the previous investigation. This capacity correlates with ability to write mirror-script with the practiced hand and with the ability to read and write inverted script. A growth factor also appears to be involved, since mirror-reading shows an inverse correlation, although not a high one, with age.

That expertness in handling spatial relationships of the symmetrical sort may be a function of the specialized anatomical organization which issues in unidextrality[2] is easily comprehensible. It is, however, a conclusion of considerable moment, particularly when considered in the light of possible variations in degree of handedness. So far as mirror-writing is concerned we were anxious to obtain more reliable data. The greatest obstacle was the difficulty in getting a satis-

[1] Downey, J. E., On the Reading and Writing of Mirror-Script, Psychol. Rev., 1914, 21, 408–441.

[2] We have adopted the suggestions of Ernest Jones concerning terminology. See Psychol. Bull., 1909, 6, 130–132.

factory index of unidextrality. We have approached the subject from several different angles. Our experiments fall into two divisions; those carried out in the spring of 1916, reported in Part I., and those carried on throughout the academic year of 1916–1917, reported in Part II.

Part I

1. *Mirror-Reading Test.*—In our experiments in the spring of 1916 we obtained a measure for ability in mirror-script reading by duplicating the method of attack reported in a previous paper.[1] We tested individually a new group of subjects, twenty-five in all, by timing the rate at which two sentences written in formal mirror-script could be read and then ranking the subjects in an order of merit (see Column 7, Table I.). The tests were each stopped at the end of eight minutes. Subjects who failed to read the sentence with such a time limit were ranked on the basis of the number of letters read correctly. The two mirror-reading tests gave by the formula

$$\rho = 1 - \frac{6\Sigma D^2}{\eta(\eta^2 - 1)}$$

a reliability coefficient of .74 (P. E. .064). This coefficient is slightly higher than that reported in the preceding paper.

This method of individual testing is somewhat tedious and later we dropped it in favor of the group method, which will be described in the second part of the paper.

For obtaining the degree of unixdetrality we employed the steadiness test, the strength of grip test, and the comparative speed and quality of handwriting of the two hands. We had planned also to use the tapping test but were prevented by circumstances from getting these records.

2. *Steadiness Test.*—As one means of estimating the relative control of the two hands we tried Whipple's test 13, using the instrument pictured on page 157 of his Manual (Part I., second edition), with stop-watch, dry battery, telegraph sounder, key, and adjustable chair. Instead of

[1] Downey, J. E., and Anderson, J. E., Form and Position in Handwriting Interpretation, *J. of Educ. Psychol.*, 1915, 6, p. 292 f.

employing the graphic method for registration of contacts, the experimenter counted them. Preliminary practice was carried to a point where we felt reasonably confident of our accuracy in counting. With regard to details of procedure, such as proper holding of the arm, placing of pencil in the hole, and the like, we conformed to the standardized instructions given in Whipple.

Since we wished to compare the relative steadiness of the two hands, a hole was selected where the subject's contacts with the supposedly inferior hand were within the range of accurate counting. As small a hole as possible was used in order to demand effort of control on the reagent's part. After the hole had been selected by preliminary trials, the subject was given the 'Ready' signal and placed his pencil in the middle of the hole; the stop-watch was then clicked. Three seconds after the signal, the circuit was closed by pressing down the key and the contacts counted from that point on for fifteen seconds. At the end of fifteen seconds, the key was released and the subject told to rest. The superior hand was first tested, then the inferior hand. Four records were taken for each hand for every subject. To obtain the relative steadiness of the two hands, the sum of contacts with the inferior hand was divided by the sum of contacts with the superior. An average was obtained of three days' records and on the basis of these averages the subjects were ranked, the subject showing the greatest difference in control in favor of the dominant hand was placed first and the one showing least difference last.

In another series of tests, steadiness was determined by the length of time the pencil could be held in the chosen hole without contact. The record was made in seconds, beginning two seconds after the 'Ready' signal. The sum of seconds in which the inferior hand could be held in position without making contact was divided by the sum of seconds in which the superior hand could be so held. The order of merit obtained by this method correlated with that obtained by the first, $+ .67$. Although the inaccuracy incident to counting of contacts was overcome by using the time-method, we

felt that, on the whole, it was less reliable than the former method because of the greater effect upon the record of a single contact. We have, accordingly, used the first steadiness test as a basis for comparison with other tests. Table I. (Column 3) gives the average record of the subjects in terms of the percentage of contacts made by the inferior hand when compared with the superior hand. The higher the percentage, the greater the difference in motor control between the two hands. One hundred per cent. would show equality in control and anything below this would indicate that the supposedly inferior hand was steadier than the dominant hand.

The utilization of the results of the steadiness test as a measure of the degree of handedness might, of course, be questioned; any one test would seem inadequate. Still more dubious might seem the arrangement on this basis of the individuals in an order of merit. Let us, however, study the records in the light of what we know of our subjects from other sources.

Subjects Wa, As and Dv are definitely left-handed, using the left hand for writing as well as for all other skilful operations. Subject Aw described himself as naturally left-handed although writing with the right hand because of outer compulsion. Our impression from conversation with many left-handed individuals leads us to believe that naturally left-handed persons resist very unequally the attempt to make them write with the conventional hand. In the same family, parental and pedagogical pressure may succeed with one child and fail with another; degree of handedness would appear operative as one factor determining degree of opposition. It seemed very possible that Aw was less left-handed than the other three left-handed subjects. A fifth subject, DJ, reported that she uses the left hand to a much greater degree than does the average person; in many situations she employs either hand with comfort.

Reference to Table I. shows that these reports of these five subjects were sustained by the outcome of the steadiness test. With one exception the other subjects appear to have

better motor control for the right hand, but in various degrees. Hi gives a better record for the left than for the right hand. Besides DJ, reagents McK, Ba and Pe show very slight superiority for either hand, as measured by motor control.

An interesting feature of our complete records concerns the matter of variability in the records from day to day. On one day each Pa, Pe, Ba, DJ and Hi gave a better record for the left hand. This aspect of the situation deserves very careful investigation, particularly in connection with its relation to the general condition of the subject. Our records are not extensive enough to justify a report.

3. *Strength of Grip.*—Strength of grip was measured by means of Smedley's dynamometer. Four trials with each hand, alternating the hands used, were given on each of two days, every effort being made to get the reagent to exert his maximal strength, with fair success, as shown by the slight variability. The percentage of strength of the inferior hand is compared with that of the superior hand. In getting this percentage we utilized the average of the eight trials. Table I. gives the results. Three of our four left-handed subjects have a stronger left than right hand; one, Wa, is stronger in the right hand. Of our right-handed group three, McG, DD and Sa show greater strength with the left hand.

Other investigators report inconsistencies in the outcome of the dynamometer experiment as an index of handedness. Mrs. Woolley from her observations on working children concludes that "superior strength of the left hand, taken alone, is no indication of left-handedness." And again, "steadiness constitutes a far better indication of left-handedness than strength, though it is by no means infallible." [1]

4. *Speed and Quality of Writing, Unpracticed Hand.*—In another series of tests we attempted a comparison of the relative skill of the group with regard to the two hands by testing writing-speed for both hands and finding the percentage number of letters written by the inferior hand when

[1] Woolley, H. T., and Fisher, C. R., Mental and Physical Measurements of Working Children, Psychol. Monog., 1914, No. 97, p. 67.

calculated on the basis of the number written by the superior or practiced hand. Starch's directions for a two-minute test of speed of writing was followed for both hands. The test was given twice and an average taken. These records are also included in Table I. The percentages vary from 29.7 to 74.8, the latter made by Aw who is naturally left-handed, but writes with the right hand. Some striking discrepancies occur with the records on the steadiness test. The speed records of the inferior hand are, however, subject to considerable correction when quality of writing is also taken into consideration.

One of us who had had considerable experience in rating handwriting attempted to arrange the left-hand specimens in an order of merit, with concealment of the names of the penmen. This arrangement was not felt to be highly satisfactory, particularly in the middle range. It was, however, quite possible to get a *group* arrangement that was fairly satisfactory. The specimens fell into seven groups. Specimens falling in groups I. and II. represent excellent left-hand writing. Table I. gives the quality of left-hand writing-quality for each subject. In a number of individual instances there is confirmation of the steadiness test. Lo, Ro, Bo, Ab, Da appear to be strongly right-handed. As is strongly left-handed; Wa and Dv are less so. Ambidextrality is indicated for DJ, Gr, OW, and DD on the basis of either speed or quality of left-hand writing. It is interesting to note that OW is father to AW; and DJ and DD are sisters.

Our method was subject to criticism in that our instructions did not discriminate carefully between speed and quality. In a repetition of this test with another group, we carefully discriminated in our instructions between the two tests, emphasizing speed in the first and quality in the second. The results were more satisfactory. We may remark in passing that we have come to expect ambidextral tendencies from penmen who write a back-slant or from those who show a tendency to shift from a running writing movement to a slow print-like type of script.

5. *Mirror-reading and Unidextrality*.—It remains to con-

sider the results of the three tests. Eighteen of our subjects were consistently right or left-handed; although with strong indications of a very great difference with respect to the extent of difference between the two hands. Seven subjects gave either inconsistent results or showed very slight difference between the two hands, namely, OW, DD, DJ, Wa, Sa, McG, Hi.

TABLE I.

Reagent	Right- or Left-Handed	Relative Unsteadiness of Inferior Hand	Relative Strength of Inferior Hand	Relative Speed Writing, Unpracticed Hand	Quality of Writing, Unpracticed Hand	Average Rank in Mirror Read	Difference in Length of Arms in 1/16 In.
Lo......	R.	350	85.5	29.7	V	2	
Wa......	L.	329.5	105.6	45.5	III	9.5	+ 1R
As......	L.	321.1	92.2	39.1	II	13	+ 8L
Bo......	R.	293.4	98.6	35.0	IV	16	+ 4R
Cl......	R.	251.5	97.3	50.8	VI	5	+ 5R
Ro......	R.	241.8	88	39.9	V	1	
Co......	R.	220.1	90.6	46.3	VI	4	
McG....	R.	217.5	124.7	35.8	V	17	
Ab......	R.	215.6	96.6	33.2	III	14.5	+ 4R
Dr......	L.	214.1	94.7	47.3	II	11.5	
Wl......	R.	210.2	59.1	41.6	III	19	
Sp......	R.	206.1	99.1	46.4	VII	14.5	+ 1R
Da......	R.	202.7	98.6	33.3	VI	7	+ 1R
Di......	R.	197.3	98.5	28.9	V	9.5	
AW.....	L.	175	89.6	74.8	II	22	+ 1.5L
OW.....	R.	155	99.4	51.7	VI	25	+ 1R
Gr......	R.	152.2	89.4	53	III	8	
DD.....	R.	146.3	107.1	36.3	I	21	+ 1R
Sa......	R.	139.9	105.8	33.4	V	3	
Pa......	R.	125.1	92.9	33.9	IV	6	+ 5R
Pe......	R.	112.1	80.3	37.9	III	23	+ 2R
Ba......	R.	106.3	85.1	34.4	IV	19	+ 2R
McK....	R.	103	97.3	36	IV	11.5	
DJ......	R.	101	93.7	57.4	II	19	+ 3R
Hi......	R.	55	95.1	37.5	V	24	0

Table I. gives the average rank of each subject for the mirror-reading tests. Lo and Ro, two of the most consistently right-handed of the group are found to excel in this test; OW, DD, DJ, McG, and Hi, five of the seven of the inconsistent group, make especially poor records.

Calculating by the formula,

$$\rho = 1 - \frac{6\Sigma D^2}{n(n^2 - 1)}$$

our coefficients of correlation for skill in mirror-reading and these various indices of dextrality, we get the following:

Mirror-Reading and Steadiness.................................. .47 (P. E. .11)
Mirror-Reading and Strength of grip............................ .03 (P. E. .14)
Mirror-Reading and speed and quality of writing unpracticed hand.. .293 (P. E. .12)
Pooling the tests for handedness by taking the algebraic sum of the deviations from the median gives a correlation with mirror-reading of... .51

The experiments with this series of twenty-five subjects give, then, some indication that skill in mirror-reading is correlated with degree of unidextrality. The evidence is cumulative and positive.

We were anxious, however, to secure better control of conditions, first in the way of a more thoroughly standardized test for mirror-reading and, secondly, in a more adequate and, if possible, simpler determination of degree of handedness. Accordingly, in 1916–1917, we approached the problem from a somewhat different standpoint.

Part II.

1. *Standardization of Mirror-Reading Test.*—In attempting to standardize the mirror-reading experiment we worked our two series of nonsense syllables, twenty-one syllables in each series (sixty-three letters), so constructed that every consonant appeared twice, once at the beginning and once at the end of each syllable and in combination with different vowels. These series of nonsense syllables, written in formal script, were etched and printed on standardized record-blanks with instructions for the test on the back of the blank. Test I. was given with a time limit of two minutes; Test II. with a time limit of one minute thirty seconds. Test II. was given immediately after Test I. Scores were calculated in terms of the number of letters correctly written in the time limit. A group of fifteen to twenty reagents could be tested at one session.

The reliability coefficient for Test I. and II. calculated from the scores of fifty-two freshmen and sophomores is .77 (P. E. .057). The correlation between these two tests is less high than expected. The instructions for the test are not wholly satisfactory. Some error arose, furthermore, from the group-testing and some from the fact that our

time limits were too long and that, consequently, in the case of subjects who finished before the signal it was necessary to record the time and calculate their score. These calculated scores might be a source of error in making an order of merit but are probably sufficiently accurate for the way in which we used them.

We tested, altogether, 180 subjects, 92 women and 88 men. Table II. shows the distribution of scores (see below).

TABLE II.

SCORE: NUMBER OF LETTERS CORRECTLY READ. MIRROR-READING I AND II

	1–9	10–19	20–29	30–39	40–49	50–59	60–69	70 and Over
I								
Men	1	6	15	17	24	14	2	9
Women	5	5	18	20	14	15	5	10
Total	6	11	33	37	38	29	7	19
II								
Men	2	6	9	5	22	22	11	11
Women	2	3	6	14	26	19	12	10
Total	4	9	15	19	48	41	23	21
Left-handed								
I	0	1	3	8	4	3	0	5
II	0	1	4	2	6	7	1	3
Right-handed								
I	6	10	30	29	34	26	7	14
II	4	8	11	17	42	34	22	18

2. *Mirror-Reading and Left-handedness.*—Having standardized the test in mirror-reading, we desired next some unequivocal method of determining degree of unidextrality. In this connection it occurred to us that studies on the inheritance of left-handedness give some indication that left-handedness is a Mendelian recessive and suggest the possibility not only that all left-handed subjects are pure in type but also that there is an "imperfection of dominance" shown in the existence of ambidexters.[1] If these suppositions were true, then on our hypothesis of a relation between skill in mirror-reading and unidextrality, a left-handed group tested for mirror-reading should show a higher central tendency

[1] Jordan, H. E., The Inheritance of Left-handedness, *Am. Breeders Mag.*, 1911, 2, 19–29 and 113–124.

than a right-handed group. Our group of 180 reagents contained twenty-four persons who claimed to be left-handed, although twelve of them write with the right hand. Table II. makes possible a comparison of the records of the left-handed and the right-handed group. There is certainly very little difference in distribution of scores. The slight difference that exists is in favor of the right-handed.

The results were disappointing but not surprising. From our previous work we had been inclined to believe that left-handed individuals show considerable variation in degree of unidextrality. For one thing, as we have mentioned before, they resist to an unequal degree the attempt to make them write with the conventional hand.

3. *Brachiometer Tests.*—While we were in the midst of this work, our attention was called to Professor W. F. Jones's method of determining congenital right and left-handedness by measurement of (1) length of the ulna, plus hand to the middle knuckle of little finger; (2) length of the humerus, (3) circumference of palm and (4) circumference of wrist. We immediately secured a brachiometer from Professor Jones together with his typed directions for taking the measures and a brief summary of his conclusions. Our information with regard to Professor Jones's work is limited to the typed pages sent out with the brachiometer and the report made by him at the 1916 joint meeting of the Psychological Association and Section L of the A. A. A. S.[1]

Very briefly, it is Professor Jones's contention that born-handedness is determinable by the measures cited above; adopted-handedness is shown by the measurement (1) of relaxed forearm circumference, (2) by contracted forearm, (3) relaxed arm (biceps) circumference, and (4) contracted arm (biceps) circumference.

The bulletin giving Professor Jones's full report upon the 10,000 measurements of paired bones made by him has not yet, we understand, been published, so that we are in the dark as to the detailed outcome of his investigation. In

[1] Jones, W. F., The Problem of Handedness in Education, Psychol. Bull. 1917, 14, p. 64.

some quarters the question will be raised whether the difference in measurement of paired bones is the cause, the effect, or a concomitant of handedness. That difference in bone length is not the result of handedness would seem to be indicated by measurements on still-born and newly-born children and by the existence of the 'shiftover' or 'transfer,' the individual who uses (inexpertly according to Jones) the shorter arm. Jones's interest in the problem is largely the pedagogical one of the discovery and treatment of the 'shiftover.'

Our interest was somewhat different; we desired a way of classifying our subjects with reference to degree of handedness. So far as we know, Jones has not been particularly interested in this aspect of the problem, namely, the amount of difference in the two arm lengths and the possible correlation of greater difference with more pronounced unidextrality —ambidextrality resulting as an accompaniment of a vanishing difference in length. Questioned on the occasion of his report before the Psychological Association as to his understanding of ambidextrality, Professor Jones answered that the 'Ambidexter' is the 'shiftover.'

Whatever the interpretation of the figures, it was highly interesting from our point of view to measure the arm lengths of our reagents. Since we found it impossible to measure all our subjects, we selected for measurement those who belonged in the best and worst quartile in the mirror-reading test. Eight of our reagents were no longer accessible. This cut our total group down to 172, with forty-three subjects in each of the two groups to be measured. Furthermore, we measured all but five of our left-handed subjects and all who were accessible of the group of twenty-five with whom we had worked in the preceding year. All together we measured some one hundred persons. Limitation of time—both on our part and that of our subjects—prevented us from taking the measurements both for born and for adopted handedness; we limited ourselves to the four bone-measurements given above, the 'ulna plus,' humerus, circumference of palm, and of wrist.

It is perhaps necessary to observe that there is a considerable margin of error in taking these measurements with the instrument used. When the difference in length between the right and left forearm is, for example, only one sixteenth of an inch a high degree of accuracy may be questioned. In our tabulation these inaccuracies would in a measure cancel one another since they would occur equally in the measurements of both groups and lie in both directions. In order to insure as great consistency as possible, all measurements, with a few exceptions, were made by the same experimenter (Mr. Payson) who, following Jones' directions carefully, developed considerable expertness in taking the measurements, each of which he verified. Another advantage in having all measures taken by one person was that a bias as to the outcome of the individual measurements was avoided since the grouping into quartiles on the basis of the mirror-reading records was made by the other experimenter. We feel, on the whole, that the measurements of the 'ulna plus' are more accurate than those of the humerus; the determination of the wrist circumference more accurate than measurement of palm circumference. It was also easier to get a satisfactory record with the men, particularly in measuring the length of humerus, than with the women.

For a majority of the persons tested the four measurements were consistently in favor of the right (or left) arm and hand; in some 39 cases inconsistencies, often very slight, were recorded. For example, the right 'ulna plus' might be longer than the left; but the left humerus longer than the right. When such inconsistencies occurred the measurements were taken a number of times; sometimes the experimenter satisfied himself as to inconsistency; at other times the difference seemed to lie within the range of error. We shall await with great interest Jones's figures on this point.

With these limitations in mind we may turn to our tabular summaries (Tables III. and IV.). We give in detail the distribution for difference in length, 'ulna plus,' for the two arms and the median and central fifty per cent. for the difference in measurement for each of the four bone-measures taken.

Table III.

Difference in Measurement Major and Minor Arm—Best and Worst Quartile (43 subjects each) Mirror-Reading I

Unit—one sixteenth of an inch for 'ulna plus'; one twentieth for wrist.

'Ulna plus' difference	0	1	2	3	4	5	6	Percentile difference
Best quartile (43)	4	6	9	9	8	3	4	.012
Worst quartile (43)	7	7	12	9	4	2	2	.009
Total (86)	11	13	21	18	12	5	6	
Wrist difference	0	1	2	3	4	5	6	
Best quartile (43)	15	8	12	3	2	1	2	
Worst quartile (43)	17	9	14	1	0	0	0	
Total (86)	32	17	26	4	2	1	2	

Table IV.

Difference in Measurements Major and Minor Arm—Best and Worst Quartile (43 subjects each). Mirror-Reading I

Units—one sixteenth and one twentieth of an inch.

	Best Quartile		Worst Quartile	
	Median	Range Central 50 Per Cent.	Median	Range Central 50 Per Cent.
'Ulna plus'	3	2–4	2	1–3
Humerus	2	1–4	1	0–4
Combined 'ulna plus' and humerus	4	3–6	3	1–5
Wrist	1	0–2	1	0–2
Palm	3	1–4	2	0–2
Cumulative measure	8	3–12	5	3–8

The difference between the two arms for the 'ulna plus' and the humerus is given in terms of one sixteenth of an inch, the unit of the brachiometer; and for wrist and palm in terms of one twentieth of an inch, the unit of the anthropometric tape. The combined measure for the 'ulna plus' and the humerus is obtained from the algebraic sum of the two measurements. The cumulative measure is simply an arbitrary index of the difference between the two arms and hands of the subjects; we took an algebraic sum of the four measures for each individual, ignoring the fact that the unit was one sixteenth of an inch in the ulna and humerus measurements and one twentieth for wrist and palm, and ignoring also the fact that the difference occurs on a much larger whole in the first two than in the second two measures.

We may meet at this point a criticism that would naturally

be made, namely, that difference in the right and left measurements should be given not in absolute terms but in terms of the relative proportion to the whole length measured. Actually, however, a variation of four-sixteenths of an inch on 17 inches and on 13 inches (extreme measurements for the 'ulna plus') shows less difference than a variation from one to four-sixteenths on either of the two given measurements. Some inaccuracy is, however, incident to giving the figures in an absolute rather than a relative way. We calculated, therefore, for the 'ulna plus' the value in percentage terms, that is, in terms of the proportional difference for each member of our group of eighty-six and obtained the average for each of our two groups; for the best quartile the difference is .012 in terms of the total length of the 'ulna plus'; for the worst quartile, .009, a slight reduction on the median as given in terms of the absolute difference. For all practical purposes the absolute difference is nearly as accurate.

A study of Tables III. and IV. indicates that with the exception of the wrist measurements, the best quartile shows a slight but consistently greater difference between the right and left hand and arm than does the worst quartile. The complete distribution of the wrist measures also shows a slight excess in favor of the best quartile. The outcome is interesting and suggestive if nothing more. Taken in conjunction with other evidence, it certainly points to a significant problem and perhaps indicates a definite conclusion.

On the other hand, a number of striking exceptions to our general conclusions occurred. Close study of individual cases would probably repay the effort required.

It is at least interesting to scrutinize certain groups in the light of the measurements. In the last column of Table I., we give the combined 'ulna plus' and humerus measurements for the fifteen subjects of the previous investigation who were available for the brachiometer measurements. So far as these measurements go they are largely confirmatory of the conclusions drawn from other tests. Reagent As, for example, whom we had on other grounds selected as the most left-handed of this group of subjects shows a more con-

siderable difference in favor of the left arm than any other left-handed subject, with one exception, that we measured. Aw and Ow (son and father) show a very slight difference in the length of the two arms, for Hi the difference is imperceptible. We have already cited the coefficient of correlation for mirror-reading and degree of unidextrality as determined by the steadiness test.

We may remark here that this group should have been more rigidly selected on the age basis, a failure which probably caused a margin of error in three or four records, particularly Bo's and Pe's, who are consistently right-handed but older than the majority of the group. The other two older reagents (OW and JD) show strong ambidextral tendencies. In Part II., also, we have ignored the age factor. But in passing, we may suggest that the relation of degree of unidextrality to age remains an outstanding problem. Various records have been given by different investigators that point to an increase of handedness at puberty. It is possible that a more extensive investigation would repay the work necessary and possibly show some interesting epochs.

Another group of subjects that we need to examine consists of the twenty-four left-handed reagents. We were able to get the measurements of nineteen of these. One of the group proved to be what Jones calls a transfer or shiftover since her right-side measurements were consistently larger than the left-side ones. She herself is very dubious as to our interpretation and says she was not shifted by an accident, so far, of course, as her knowledge goes, and has no inclination to be right-handed. Her mother is left-handed and it is possible that the left-handedness of the daughter may be explained by imitation.

The other eighteen left-handed persons show considerable variation in the amount of difference between the two arms. It varies from 0 to 9, for the combined arm length. The extreme measurement is less than our extreme in the right-handed group but this would be anticipated from its smaller number. Eight of the eighteen write with the minor hand, in this case the right hand. With two exceptions this group

of eight shows only a slight difference in the measurements of the two arms, and, in general, a less difference than that found for most of the subjects of the other group. The exceptions in the first group are its two oldest members. Parental and pedagogical pressure in forcing a shift from the left to the right hand in writing is becoming milder. Undoubtedly under present school-room conditions these two individuals would be writing with the left hand.

About four out of a hundred individuals are born left-handed according to Jones but three of them are using the minor hand. Among the one hundred supposedly-right-handed persons whom we measured we found three who gave consistently for the four measures an excess in favor of the left side. One of the subjects reported himself as ambidextrous, and in fact the difference in favor of the left is very slight. He is very inexpert at mirror-reading. The other two belong in our best quartile for mirror-reading. One of them admitted that he was left-handed. He calls himself right-handed because of writing with the right hand; he does, however, a number of things with the left hand. No particular pressure was exerted to make him write with the minor hand. As a matter of fact the difference between the two hands is not extreme. The third case shows an extreme difference between the two arms and hands. Questioned in detail it is obvious that this subject does many things with his left hand. He supposed this was due to accident in forming certain habits. After his attention was called to the matter he reported a considerable number of situations in which he used the left hand in preference to the right.

Besides these three reagents seven others gave an excess of difference in favor of the left arm if we take the combined measures of the 'ulna plus' and humerus; their records, that is, are inconsistent. The difference, however, is very slight except in two instances. Five of the seven are very inexpert mirror-readers. This group would be a most interesting one to study.

4. *Puzzle-tests.*—Throughout our work on mirror-reading the impression has persisted that facility in such reading is

merely a special instance of skill in handling spatial relationships. The puzzle-tests to be described in this section of the paper were designed for further investigation of the nature of this ability. It seemed probable that one of two things was true. Either unidextrality was connected with a general ability to appreciate all sorts of form relationship or it was confined to an understanding of symmetry. Two kinds of puzzles were therefore designed, one of them laying emphasis on the symmetrical relationship and the other requiring, instead, an aptitude for the manipulation of form. Three puzzles of each kind were used. Our subjects were twenty-five freshmen.

Series I. Symmetrical Puzzles.—The puzzles in this group vary somewhat in appearance and character but exhibit in the main the same fundamental principles. To each of these puzzles there were two parts; a diagram and a dissected symmetrical counterpart of the diagram. The pieces of the diagram were numbered and the pieces of the symmetrical part removed and placed on the table beside the subject. The subject was required to place the pieces *in the order in which they were numbered on the diagram.* To do this it was necessary for the subject to refer constantly to the diagram and make the translation into its symmetrical form. It was thought that this ability was very similar to the ability to translate symmetrical writing, that is, mirror script. The following directions were given before the puzzles were shown to the subject.

"Do you understand the difference between symmetrical and similar figures? Symmetrical figures have essentially the same form but the parts are reversed in such a way that they may not be fitted together. Your two hands, for example, are symmetrical.

"I am going to give you some puzzles to fit together that involve the principle of symmetrical figures. You will have a diagram to guide you which is symmetrical to the figure you will fit together and so you will have to keep the difference between symmetrical and similar figures always in mind or you will become confused.

"The pieces are numbered on the diagram in the order in which they are to be placed. Do not turn them over."

When it was certain that the subject understood the directions, the puzzles were produced. The parts were placed on the table at one side of the puzzle in a given order, so chosen as to give the subject no hint as to the solution. The time required to place each piece was taken and an order of merit obtained based on the time required to do the entire puzzle.

As previously mentioned the symmetrical puzzles differ somewhat from one another and a brief description of each is given below.

Symmetrical Puzzle No. 1.—This puzzle is rectangular in form and is composed of six differently shaped figures bounded by irregular curves and straight lines. It is, perhaps, the easiest to handle of the three. The pieces are numbered in an irregular order.

Symmetrical Puzzle No. 2.—This and the third puzzle are rather similar in design. Both are circular in outline and both composed of parts radiating from a central point. Puzzle No. 2 is formed of rather similar parts numbered consecutively. In this puzzle it is particularly difficult to recognize the separate parts from inspecting the symmetrical diagram. Anyone not possessing a keen perception of form might be expected to have difficulty with this puzzle.

Symmetrical Puzzle No. 3.—This puzzle differs principally from the preceding one in having parts that are different enough from one another to be more easily recognized by aid of the symmetrical diagram. It is probably less a test of form perception. The pieces are numbered in an irregular order.

The method pursued in these tests—which involved the placing of the dissected pieces in the order indicated by the number on the diagram—eliminated one element of variability in individual manipulation of puzzle pieces and made it possible to record the time needed for placing each insert, a necessary preliminary for standardization of such a test. It did not, however, overcome blocking and possibly increased

it since free trial and error was in a certain measure ruled out. A possible improvement in method might be made by limiting the time for placing of each insert. Some method of quantitative standardization seems necessary if we are to utilize puzzle tests in any other than a purely individual way. Many authorities assert, however, that the method of handling a puzzle is throughout much more significant than is the quantitative factor.

Series II. Angular Puzzles.—The first puzzle of this series is that known as the Kempf-Knox Diagonal Test. The two others are somewhat similar to it in design but more difficult. They are composed each of six triangular pieces and are so constructed that three of the pieces, if properly placed, form one half of the puzzle exactly symmetrical to the other half. It was thought that a good subject would realize that the pieces were in pairs, solve half of the puzzle and then easily construct its symmetrical half. As a matter of fact very few, if any, of the subjects made use of this line of reasoning; consequently, the puzzles were as difficult as if they had been composed of six dissimilar pieces. The instructions given were extremely simple. The pieces were spread out upon the table and the subject told to fit them into the frame. The only record kept was the total time required to complete each puzzle.

In criticism of these angular puzzle-tests it may be said that they proved to be too difficult. A series as simple as the Kempf-Knox Diagonal Test would be preferable. Blocking on a single piece may interfere with the record in such a way that a poor subject happening to find a good combination may make a much better record that a good subject not so fortunate in the placing of the first pieces. With all these imperfections in method, individual differences are plainly evident. Some subjects are very evidently good at manipulating the blocks so as to fill a given space while others are just as evidently confused.

To return to our original question. An order of merit based on the average time required for solving the symmetrical puzzles correlated with that based on two tests for

mirror reading .43 (P. E. .11). An order of merit based on the average of the Angular Puzzles No. 2 and 3 gave a correlation of .15 (P. E. .13). These figures offer some evidence of relationship between the mirror-reading test and the solving of symmetrical puzzles but are inconclusive so far as angular puzzles are concerned. Many expert mirror-readers are so obviously superior in their fitting together of puzzles that we feel that investigators of puzzles tests would do well to keep in mind unidextrality as a possible factor in the case.[1] No doubt, close study of our individual cases that represent discrepant results would yield valuable material.

5. *Mirror-Reading and General Intelligence.*—Mirror-reading is, we believe, a highly specialized capacity. It is a common occurrence to find surprising expertness or inexpertness in unexpected quarters. As a side-light on this aspect of the situation it is interesting to cite the correlational coefficient between an order of merit based on skill in mirror-reading and one based on the Intelligence Quotient obtained by use of the Stanford Adult Tests. I. Q's were at hand for forty of the subjects tested in mirror-reading; the correlational coefficient was .14 (P. E. 11). This coefficient does not, of course, prove that general intelligence is of no value in adaptation to the test. With a small group such a coefficient is indecisive but at least it weighs against the possibility of a very high correlation. The outcome is, moreover, in agreement with the failure to find any indication of a correlation between mirror-reading and the completion test, as reported in a former paper.[2]

6. *Mirror-Reading and Visual Capacity.*—Throughout our work on mirror-reading we have had a strong feeling that some connection exists between this and visual capacity (visual imaging, visual attention and the like). Certainly we are collecting an increasing number of clever abstract thinkers, often literary in type of mind, who are hopelessly

[1] The twelve superior subjects in both sets of puzzles gave an average of arm-measurement in excess of that of the twelve inferior subjects. But the subjects in each group are so few and the M. V. so large that the average has only suggestive value.

[2] Downey and Anderson, *loc. cit.*, p. 360.

adrift when confronted with mirror script, and of the slower visual type who handle it with perfect ease. The coincidences seem too frequent to be accidental. It is, of course, quite possible for one's observations to become biased; and in the absence of an objective test for visualization one cannot discuss the problem with any satisfaction. We attempted a study of two small groups of freshmen, six in each, chosen from the two extremes in mirror-reading and exhibiting in the one case considerable difference in the bone-measures and very slight difference in the other. Four of the latter, or poor, group are very deficient in visual control as evidenced by their scores and their introspective reports in the clock test (Stanford revision, XIV. 6) Code Test (Stanford revision, XVI. 6) and the Cube Test (Langfeld and Allport)[1] but two of this group seem to handle visual imagery with ease. Of the good group four give evidence of excellent control over their visual imagery, one is a fair visualizer, one is somewhat deficient. The relationship is one that deserves detailed investigation not only because of the light it may throw on the explanation of mirror-reading but also because of its possible complication with unidextrality.[2]

7. *Summary and Conclusions.*—As an outcome of our work we are convinced that the proposition that asserts a relationship existing between mirror-reading capacity and unidextrality deserves serious consideration. Furthermore, skill in mirror-reading appears to be indicative of a specialized capacity which is of value in some forms of spatial orientation. There is also a possibility that mirror-reading and visual capacity are related abilities. The real point of interest in all this is not, of course, the citing of factors concerned in mirror-reading as mirror-reading. *This is incidental to a bigger question whether or not unidextrality, or rather degree of unidextrality, is bound up with certain mental attributes that, taken together, constitute a somewhat definite mental type.* In

[1] 'An Elementary Laboratory Course in Psychology,' p. 72.

[2] We have at hand some evidence of correlation between speed of mirror-reading and speed in reversed spelling of words (a visual test). The latter test *if standardized* would be of considerable value.

the absence of evidence suggestions must be recognized as purely speculative.

The real antithesis appears to lie between ambidextrality and unidextrality and not between right- and left-handedness. Possibly there are two factors involved, one governing the degree of dominance and the other placing it. Close study of family groups should be instituted. Quite likely we should find the very right-handed and the very left-handed persons in the same family.[1] To make our records of value we must have some way of checking up reports. If large groups are to be studied this method must be a comparatively simple one. It is for this reason that Jones's measurements seem to us so promising, and that quite apart from the questions concerning the causal or indirect relationship existing between handedness and difference in length of paired bones. At present we are pretty much in the dark as to the anatomical basis of unidextrality. In particular, we should like to know more of the cerebral conditions found to exist in the case of individuals of pronouncedly unidextral or ambidextral tendencies. Here we must await neurological investigation.

Meanwhile, there are a number of investigations that need reviewing with the possibility in mind that approximate ambidextrality and extreme unidextrality may be correlated with somewhat definite mental differences. One of these reviews should consider the evidence that seeks to relate speech defects to shifting from the natural hand in writing. Certainly we have need for more precise data both with reference to the kind and extent of speech difficulty and the degree to which a dominant-handedness actually existed. Again, evidence that mental defectives are predominatingly left-handed or ambidextrous should be reviewed in connection with more adequate measurements of what constitutes dextrality, sinstrality or ambidextrality. Doll,[2] in harmony with other investigators, observes that "sinistrality in grip does not necessarily imply sinistrality in other manual activities"

[1] Should extreme unidextrality be found correlated with a determined mental type, we should have in an index of unidextrality a measure of more than handedness, and the way would be open for enlightening studies on human inheritance.

[2] 'Anthropometry as an Aid to Mental Diagnosis,' p. 65.

and yet often the judgment as to the dominant hand is based upon the relatively stronger grip. In particular, it might be worth determining whether mental defectives show striking variations on the manual side and on the verbal that might possibly serve to place them in groups. If so, the possible unidextrality of the first (or objective-minded) and the approximate ambidextrality of the second (or subjective-minded) group would be an interesting question for investigation.

THE DIFFERENTIAL SPATIAL LIMEN FOR FINGER SPAN

BY HERBERT SIDNEY LANGFELD

Harvard University

The investigation of the limen for finger span was suggested by the practical demand for accuracy in estimating the number of objects or size of objects grasped between the thumb and index finger. The problem of psychophysical measurements in general, however, although one of the oldest, has rather increased than diminished in theoretical interest and importance, as the many functions involved, and at first so little suspected, have come to light, thus casting suspicion upon the accuracy of threshold measurements. In representing these results, therefore, the theoretical question of method and the desire to present further data for discussion are uppermost.

Beside the establishment of the limen, the experiment was arranged to ascertain individual differences, differences according to the method of presentation of the stimuli, the relation of the subjective feeling of certainty to the objective accuracy, and the value of the doubtful or guess judgments.

There were six subjects—five men and one woman. Three of them, *A*, *B*, and *F*, were trained experimentalists, of which *B* was the best trained. The other subjects had little knowledge in methods. Subjects *A* and *B* were slow, deliberate, and careful. The others were quicker and of a more nervous temperament. Subject *E* was particularly high strung and nervous to the extent of bodily twitchings. This description was corroborated by another experimenter who had all but one of these subjects.

The instruments used were six calipers, two of which were Starrett's No. 25 M as shown in the cut, page 419. The other four were modelled from them. Small, accurately machined right-angle plates with perpendicular surfaces,

16 mm. by 25.5 mm., were securely screwed on the ends of the two points (*A*). At the end of the adjustable arm rest were upright blocks which held the clips (*B*) into which the calipers could be easily slipped. The arm rest was adjusted for each subject so that when the hand rested comfortably on the board with fingers extended over the edge, the caliper plates could be grasped in their middle without moving the hand.

The procedure adopted was the method of constant stimuli. Not only did this method seem less open to theoretical criticism, but it was more practical for this experiment than the method of limits, since it would have been difficult to obtain an instrument that could be quickly adjusted with the necessary degree of accuracy. The standard distance was 5 cm., a distance arbitrarily selected as not being too great to be comfortably spanned by all the subjects and yet sufficiently large that relatively large differences could be used. The comparison stimulus was set at a distance of 1 mm. Five comparison stimuli were used, the two extremes being eliminated.[1] There were, then, two standard stimuli of 5 cm. each and four comparison stimuli of 5.2 cm., 5.1 cm., 4.9 cm., and 4.8 cm. respectively.

A long series of about seven hundred judgments for each subject was run through during the first half year and the effect of practice was thus in great part eliminated.[2] The results have not been included in this paper, and the method employed, which differed from that finally adopted, need not be described here.

The main experiment followed three weeks after this practice series (an unfortunate break due to the mid-year period). Three separate limens were obtained: one for the span of the index finger and thumb of the right hand, the standard and comparison stimuli being presented successively

[1] See Fernberger, S. W., On the Elimination of the Two Extreme Intensities of the Comparison Stimuli in the Method of Constant Stimuli, Psychol. Rev., 1914, **21**, 335–355.

[2] Fernberger, S. W., The Effects of Practice in its Initial Stages in Lifted Weight Experiments and its Bearing Upon Anthropometric Measurements, *Amer. J. of Psychol.*, 1916, **27**, 261–272.

to the same hand; another for the two hands, the stimuli being presented simultaneously to the two hands, the standard to one and the comparison to the other; a third, also for the two hands, but with the stimuli presented successively, first the standard to the one hand followed by the comparison to the other, and vice versa.

The subjects worked in pairs, alternating after each set of thirty judgments in order to prevent fatigue, and only ninety judgments were made in the hour. A week intervened between each group of ninety judgments. Each set of thirty judgments consisted of ten judgments for each of the three limens. The subject made ten judgments for the limen of the right hand, then ten when the stimuli were presented simultaneously to the two hands, and finally ten when the stimuli were presented successively to the two hands. On half the days this order was reversed, the stimuli being presented first to the two hands successively, and last to the right hand alone. The subject then read while the other subject went through the thirty judgments. The series of ninety judgments for the two subjects lasted about an hour. The entire experiment extended through the four months of the second semester.

The order of the presentation of the comparison stimuli was 5.2 cm., 4.9 cm., 5 cm., 5.1 cm., and 4.8 cm., with the comparison stimuli presented first in the successive presentation and to the right hand when two hands were used. This same order for the comparison stimuli was then repeated with the standard first and to the left hand when two hands were used, the time and space error being thus objectively eliminated. The judgment was given in terms of the second stimulus when successive and in terms of the right hand stimulus when simultaneous. The stimuli were applied at intervals of a second.

The subject sat at a convenient height in front of the table with his hand lying comfortably upon the rest, which was properly adjusted so that the fingers could grasp the plates at their centers. All the calipers were so set that the two plates were equidistant from the middle point of the

instrument. In thus keeping the center of the span always in the same place, there was no possibility of detecting a difference through a change in the symmetrical relation of the two plates to the hand. The subject closed the second, third, and fourth fingers and extended the thumb and index fingers, which were held a convenient distance apart. He closed his eyes, the experimenter placed the calipers and clips, and said 'Now.' The subject then closed his fingers on the two plates and gave his judgment. He was never at any time given a clue as to the accuracy of the results.

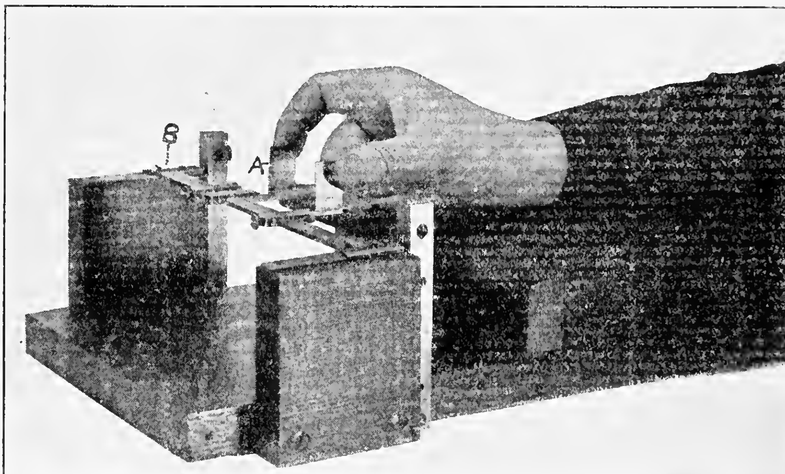


Fig. 1. Arm Rest with Caliper and Hand in Position.

The subject was instructed to clasp the plate with sufficient pressure to be able to feel the spatial extent separating them, the sensations involved being principally those of joint and tendon. He was particularly cautioned not to use the distance moved through by the fingers in clasping as a criterion. Special attention was paid to practice in abstracting from this movement perception during the five months of preliminary training, and from the introspection of all the subjects, it is fair to assume that there was the proper direction of attention and that the judgment was based on the static spatial perception.[1] The subjects were told to judge

[1] Inasmuch as the fingers did not always assume the same position previous to the grasping, this criterion of movement would have been a very uncertain one.

whether the second space was larger, smaller, or equal to the first, or in case of simultaneous presentation, whether the right was larger, smaller, or equal to the left. He was further told that there were objectively equal stimuli. He was directed to say equal only when there was an actual perception of equality, and not when there was a mere perception of no difference. He was directed to grade his judgments *a*, *b*, and *c*, according to his degree of confidence in his answer, *a* being the highest degree. If he could not make a larger, smaller, or equal judgment even with *c* confidence, he was to say 'equal *d*,' and was then to guess either larger or smaller. Under this category would fall all doubtful judgments, no difference judgments, and or-judgments.[1] Both judgments were noted for separate treatment. The response under these instructions, for example, took the form of 'larger *c*,' 'smaller *a*,' 'larger *b*,' etc. None of the subjects had the least difficulty in grading their confidence, the confidence response coming immediately and without hesitation after the first judgment. Even during the first series, no one complained of it interfering with his judgment, and as it did not take an appreciably longer time, it was decided to continue it throughout the series even though correlations of degree of confidence with accuracy of results and an analysis of the guess judgment had been made by previous investigators.

The method of treatment of the results was according to the $\varphi(\gamma)$ hypothesis.[2] The upper and lower limens, S_2 and S_1, the interval of uncertainty, $S_2 - S_1$, the Measurement of Sensitivity or Threshold of Volkmann, $(s_2 - s_1)/2$, subjective equality, $(c_2 + c_1)/(h_2 + h_1)$, and coefficients of precision, h_1 and h_2, were calculated. The results are recorded in Table I.

Directing our attention first to the M. S.[3] for the one hand,

[1] S. S. George believes that doubtful and or-judgments probably belong to the same category. For an excellent evaluation of these various judgments and an explanation of their non-serial nature, see his article entitled Attitude in Relation to the Psychophysical Judgment, *Amer. J. of Psychol.*, 1917, 28, 1–37.

[2] See F. M. Urban, The Method of Constant Stimuli and Its Generalization, Psychol. Rev., 1910, 17, 229–259. Also, S. W. Fernberger, On the Relation of the Methods of Just Perceptible Differences and Constant Stimuli, Psychol. Monog., 1913, No. 61, especially pp. 29 and 38. I am greatly indebted to Professor Fernberger for his valuable advice in regard to the mathematical treatment.

[3] 'Measurement of Sensitivity.'

it will be noticed that there is a considerable difference in the individual M. S. as well as in the coefficients of precision. The two subjects who have the highest coefficients of precision, subjects *A* and *B*, are two of the three best trained

TABLE I

Method	Subjects	h_1	h_2	$S_1=\left(\frac{C_1}{h_1}\right)$	$S_2=\left(\frac{C_2}{h_2}\right)$	S_2-S_1	$\frac{S_2-S_1}{2}$	$\frac{C_2+C_1}{h_2+h_1}$
Right hand	*A*.......	6.1604	5.6587	4.976	5.066	0.09	0.045	5.019
	B.......	5.520	5.042	4.94	5.03	0.09	0.045	4.987
		(4.9632)	(4.849)	(4.917)	(5.049)	(0.132)	(0.066)	(4.982)
	C.......	2.5922	3.7562	4.977	5.075	0.098	0.049	5.035
	D.......	3.5630	3.3988	5.009	5.074	0.065	0.032	5.044
	E.......	3.1709	2.9872	4.978	5.042	0.064	0.032	5.009
	F.......	(4.2877)	(4.5568)	(4.98)	(5.056)	(0.076)	(0.038)	(5.019)
Simultaneous	*A*.......	3.5461	3.978	4.982	5.081	0.099	0.049	5.035
	B.......	3.1252	2.809	4.822	5.097	0.275	0.137	4.952
		(2.9785)	(2.4428)	(4.772)	(5.135)	(0.363)	(0.181)	(4.935)
	C.......	1.3948	2.5185	4.966	5.002	0.036	0.018	4.989
	D.......	2.0402	2.5045	4.929	5.15	0.221	0.11	5.051
	F.......	3.9459	4.2674	5.025	5.035	0.01	0.005	5.03
	F.......	(1.6789)	(2.868)	(4.736)	(5.096)	(0.36)	(0.18)	(4.963)
Successive	*A*.......	4.0334	4.7442	4.97	5.087	0.117	0.058	5.033
	B.......	3.1591	2.412	4.894	5.102	0.208	0.104	4.984
		(2.4178)	(2.1034)	(4.821)	(5.155)	(0.334)	(0.167)	(4.977)
	C.......	2.5538	3.3941	4.933	5.047	0.113	0.056	5.002
	D.......	1.7494	2.6402	5.007	5.137	0.131	0.065	5.085
	E.......	2.6394	3.0165	5.02	5.068	0.048	0.024	5.046
	F.......	(2.0141)	(2.7383)	(4.951)	(5.058)	(0.107)	(0.053)	(5.012)

h_1 = Coefficient of precision "less judgments,"
h_2 = Coefficient of precision "greater judgments,"
S_1 = Lower limen,
S_2 = Upper limen,
$S_2 - S_1$ = Interval of uncertainty,
$\frac{S_2 - S_1}{2}$ = Measurement of sensitivity (Threshold of Volkmann),
$\frac{C_2 + C_1}{h_2 + h_1}$ = Point of subjective equality.

subjects, and the two who were the most careful and consistent in their attitude toward the experiment. Even in the other two methods, subject *A*'s coefficients are the highest, and *B*'s coefficients rank very high. Although some of the other subjects had lower M. S.'s, it seems preferable to consider the M. S.'s of these two subjects, which are identical, as the M. S. under the most constant conditions, rather than to average all the subjects. The data from the other subjects have been included, however, for comparison. The

M. S. is, then, 0.45 mm., which bears a ratio to the standard of about 1/100.[1]

The data of subject *F* have unfortunately little value and have therefore been bracketed. This subject and subject *B* had a great many doubtful judgments. For subject *B* there are available a sufficient number of judgments beyond the fifty required to permit the elimination of these doubtful judgments entirely and to substitute other judgments in their place. This was impossible in the case of subject *F*, so that the data include the doubtful judgments as equal. The data were calculated for subject *B*, both with and without the doubtful judgments included, the calculations for the former being included (bracketed) for comparison. It will be seen that with the doubtful judgments eliminated, the M. S. is lower and the precision higher. An explanation for this is given on page 427 where the doubtful judgments receive further treatment. Another factor that makes subject *F*'s results questionable is that during work in another university he had acquired a set not to give equal judgments and it was rather difficult for him to change this.[2] That may be a reason why he had so many doubtful judgments. His data are only used in discussing relative values. The other five subjects had very few doubtful judgments, so few that it was not necessary to eliminate them, as they made no appreciable difference in the results. The exact number of doubtful judgments for the subjects in all three methods, that is in seven hundred and fifty judgments, is as follows: subject *A*, 3; *B*, 100; *C*, 10; *D*, 14; *E*, 6; and *F*, 86.

The results of the subjective equality increase the confidence in the method. In all three methods of presentation and for all subjects, it approaches closely to objective equality. When we compare the results of the three methods of presentation, we find little consistency among subjects. From experience we would have prophesied higher M. S.'s and less

[1] Jastrow some years ago made some experiments upon finger span with wooden blocks of various sizes. He found that with a difference of 1/100 the subject made 3.5 errors in ten judgments. The Perception of Space by Disparate Senses, *Mind*, 1886, p. 552.

[2] See S. W. Fernberger, The Effect of the Attitude of the Subject upon the Measure of Sensitivity, *Amer. J. of Psychol.*, 1914, 25, 538–543.

precision for two hands. For precision this is the case, with the exception of subject *E* for simultaneous presentation. As regards the M. S., the subjects vary among themselves, even the two whom we have selected as most reliable. Subject *A*'s M. S. for the right hand is the lowest, but it is only slightly lower than his simultaneous, and not much lower than his successive M. S. Subject *B*'s M. S.'s, on the other hand, are much higher for simultaneous and successive presentations. Four subjects' M. S.'s are lower for the right hand, and two for simultaneous presentation. Some are lower for simultaneous than for successive presentation, and others the reverse. There are so many possible factors involved that it is impossible from the data at hand to offer satisfactory explanations. There is no doubt, especially with untrained subjects, that various criteria were used. Judging from subject *A*'s consistent results, it is probable that his attitude and his criteria of judgment remained fairly constant. But from what has recently been found in regard to the effect of attitude on limen,[1] it is useless to discuss individual differences in regard to sensitivity, unless one is sure that the attitude is constant or is cognizant of the nature of the change.

The ranking of the five subjects according to their M. S. for the three methods is:

Right hand....................................	*E D* (*A B*) *C*
Simultaneous..................................	*E C A D B*
Successive....................................	*E C A D B*

There is perfect correlation between the two methods for the two hands. The change of rank appears in the change from one hand to two hands, the degree of correlation being only .18.

The exceedingly low, in fact very improbable M. S. of .01 mm. of subject *E*, makes one suspect that in his case fifty judgments for each pair were not sufficient. It seems too low even though his nervous jerky manner in holding the stimuli may have changed the simultaneous presentation into a quick and repeated alternation of stimuli.[2]

[1] See Fernberger, *op. cit.*, and George, *op. cit.*

[2] While on the subject of the number of judgments, it might be said that the decrease of all the coefficients of precision for the two-hand judgments is an indication that it would have been better to have made more judgments for these series.

Another point of interest is that the S_2 is further from the objective equality than the S_1, in practically all cases. This indicates the constant error or general tendency of judgment and the shifting here is in the direction of the larger judgments. If more subjects had been used, it is possible that there would have been those who preferred smaller judgments and those who showed no preference, that is, negative and indifferent types as well as positive types, such as Martin and Müller found with lifted weights.[1] There are certainly not sufficient cases to conclude that only the positive type exists in determining this limen, although this type seemed to predominate in most of the previous experiments. On the other hand, one of the explanations offered relative to the lifted weight experiment, namely, that the types depend upon the strength of the subject, is hardly applicable here, unless it happens that one develops a general set in the course of judgments depending primarily upon strength and that then this set influences all judgments of greater or less.

In Table II. are recorded in percentage all *a*, *b*, and *c* confidence according to right and wrong judgments. The figures are calculated from the data of all six subjects. When the $d = 0$, there is little difference in the per cent. of *a* judgments for the larger and smaller judgments. The figures are 21 per cent. compared to 17 per cent. But when $d = 1$ mm., many more correct answers have *a* confidence than incorrect, 20 per cent. and 25 per cent. correct to 8 per cent. and 9 per cent. incorrect. When $d = 2$ mm., this difference is still greater, 31 per cent. and 34 per cent. correct to 10 per cent. and 9 per cent. incorrect. The distribution of the *b* judgments is in the same direction, but the differences are not so marked.[2] It is evident that the mere judgments

[1] 'Zur Analyse d. Unterschiedsempfindlichkeit,' Leipzig, 1889. Fullerton and Cattell among others found that the judgments were more often right when the second weight was heavier. 'On the Perception of Small Differences,' University of Pennsylvania Publications, Philosophical Series, 2, 1892, p. 127. It seems impossible entirely to eliminate the time and space errors and obtain symmetrical curves. If other investigators had not found a similar lack of symmetry we should have thought we had perhaps too few judgments. Even so, fifty judgments seem the minimum and one hundred would be better.

[2] Pearse and Jastrow found a similar relation between confidence and correctness.

'greater' or 'less' do not give a complete picture of the effect of the difference. If we had used categories such as 'much greater' and 'much less,' probably the differences which here were expressed in a subjective scale would be expressed in the actual judgments. Suppose 'greater' was the correct judgment, then if conditions were the same as in this experiment, there would have been more 'much greater' judgments than 'much less' judgments, if these categories had been added.[1]

Table II

Com. Stim.	*a*	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*
		<			=			>	
5 $d = 0$	66	111	201	0	16	130	71	106	156
	17	30	53	0	10	90	21	32	47
		< (wrong)			=			> (correct)	
5.2 $d = 2$	12	44	63	0	8	91	198	225	223
	10	37	53	0	8	91	31	35	34
		> (wrong)			=			< (correct)	
4.8 $d = 2$	8	19	62	0	19	115	212	191	221
	9	21	70	0	14	86	34	31	35
		< (wrong)			=			> (correct)	
5.1 $d = 1$	16	73	115	1	23	118	105	180	217
	8	36	56	0	17	83	25	25	50
		> (wrong)			=			< (correct)	
4.9 $d = 1$	15	48	101	0	14	145	106	185	241
	9	30	61	0	9	91	20	35	45

a, *b*, *c* = degrees of confidence. The first horizontal column of figures for each comparative stimulus shows the number of judgments for each degree of confidence and for each class of judgments. The second horizontal column shows the per cent. of each of the three degrees of confidence for each class of judgments.

See Jastrow's A Critique of Psychophysical Methods, *Amer. J. of Psychol.*, 1887–1888, p. 305. Fullerton and Cattell also found that the degree of confidence varies as the percentage of right guesses. *Op. cit.*, p. 126. Harold Griffing, in his article on Sensations from Pressure and Impact, Psychol. Rev. Monog., 1895, p. 45, reports the same relation. As there were very few *a* judgments, this relation holds for the *b* judgments. G. M. Whipple, in An Analytic Study of the Memory Judgment and the Process of Judgment in the Discrimination of Clangs and Tones, *Amer. J. of Psychol.*, 1901, p. 446, also found much greater certainty for right answers.

[1] H. Ebbinghaus advised the use of the following categories: deutlich kleiner, ebenmerklich kleiner, gleich, ebenmerklich grösser, und deutlich grösser, 'Grundzüge der Psychologie,' Bd. 1, 1892, p. 74.

It is also interesting to note that the equality judgments were never given with confidence and the amount of *b* confidence is less than with the other judgments.

The subjects, ranked in order of confidence according to the number of *a* judgments, are: subject *E*, 311; subject *A*, 280; subject *D*, 129; subject *B*, 38; subject *C*, 37; subject *F*, 15 (see Table III.). It is seen that only subjects *E* and *A*

TABLE III

DISTRIBUTION OF THE JUDGMENTS OF EACH SUBJECT ACCORDING TO CONFIDENCE

Confidence	Subject *A*	Subject *B*	Subject *C*	Subject *D*	Subject *E*	Subject *F*
a..........	280	38	37	129	311	15
b..........	168	170	230	270	273	174
c..........	298	448	472	334	162	478

had a high degree of confidence. Subject *D* had a medium degree, and subjects *B*, *C*, and *F* very little. This assumes that they are using the same scale. Even though one subject's *a* means a higher degree of confidence than another's, however, it seems as if the number of *a*'s should give some indication of the subjective state. There is the usual lack of agreement between accuracy and confidence. Subject *A*'s results do correlate positively, but subject *B*'s correlate negatively. On the other hand the correlation between the M. S. for the right hand and degree of confidence is high, being .83. The correlation between the M. S. for the other two forms of presentation, simultaneous and successive, and the degree of confidence, however, is only .40. As one would suppose, we do find a certain correlation between lack of confidence and the number of doubtful judgments. Subjects *B* and *F* have the most doubtful judgments. They also are low in confidence. On the other hand, because a subject lacks confidence is no indication that he will have many doubtful judgments, as is shown in the case of subject *C*, who has only ten doubtful judgments and only thirty-seven *a* judgments out of seven hundred and fifty.

It will be recalled that the subjects were instructed that when they could not answer 'larger,' or 'smaller,' or 'equal,' they were to record the fact by naming it a *d* judgment and

then guessing either larger or smaller. There was found to be among all the judgments made in the half-year, including the preliminary trials of the main series, two hundred and twenty-nine such guesses when there was an objective difference. Of these, one hundred and sixty, or 70 per cent., were correct, and sixty-nine, or 30 per cent., incorrect. There were forty-nine guess judgments with objective equality, namely, when $d = 0$. Here, in contrast with the above distribution, there are almost an equal number of larger and smaller judgments, namely, twenty-six and twenty-three. These results conform to the laws of chance, although there are too few cases to be at all convincing, except in connection with the other results. Most of the guess judgments were divided between subjects *B* and *F*. Subject *B* had 72 per cent. correct and 28 per cent. incorrect; subject *F*, 71 per cent. correct and 29 per cent. incorrect. All the others showed the same tendency, but the numbers were too small to have any value except to strengthen the evidence of a predominance of correct answers by leaning in the same direction. It is clear, therefore, that although the subjective state of the subject was one of extreme doubt with no conscious attitude to swing him one way or the other, if he made a judgment without deliberation or hesitancy, he would have had many more correct answers than incorrect.

In thus making 'greater' or 'less' judgments, the subject probably assumes an attitude more closely resembling his set toward the other judgments. The 'doubtful set' is undoubtedly a change of attitude toward the experiment and should not be encouraged. It seems as if the instruction to guess 'larger' or 'smaller' gradually brings the subject into a constant attitude or to one approximating it. This seems to be borne out by the decrease in doubtful judgments. George says that ". . . the observers tended in the effort to maintain a constant receptive attitude to make passive judgments." Paradoxical as it may seem, the attitude when one guesses in the above situation is more nearly like the passive attitude as contrasted with the reflective, deliberate attitude of the doubtful judgment. George says further

that "the indication is that the way to maintain a constant attitude is to dispose oneself to make passive immediate judgments."[1] Our experience seems to support this. One's attention is diffused, as it were, over the experience, rather than concentrated now on this, now on that, criterion, and the response comes, it often seems, almost involuntarily. One is often surprised that the judgment is correct. There seems to be little doubt that the subjects of the psychophysical experiment should be trained in this attitude.[2]

In this connection, subject *B*'s two sets of results are of value. The ones in brackets, in which the doubtful judgments have been treated as equal judgments, give higher M. S.'s and lower precision than the results in which these doubtful judgments have been ruled out and judgment of *a*, *b*, or *c* certainty substituted. From the relation of the right to wrong judgments for this set, it is evident that if, instead of ruling out the 'guess' judgments and substituting others, the 'larger' and 'smaller' guess judgments had been used, the results would have been similar to those obtained from the former procedure and which have been used in the table unbracketed.

R. S. Woodworth recently reopened the question of the use of the equal and doubtful category.[3] He says on page 66 that "the main point is, of course, that the middle class of judgments, consisting of 'equal' and 'doubtful' judgments, is only a subjective definition, varying from individual to

[1] *Op. cit.*, p. 33.

[2] Several investigators have analyzed guess judgments and our results are in agreement with theirs. Pearce and Jastrow, On Small Differences of Sensation, *Memoirs of the National Academy of Sciences*, Vol. 3, 1884, found 60 per cent. correct guesses. Fullerton and Cattell found 60 per cent. and 65 per cent. correct judgments for two subjects. *Op. cit.*, p. 132. Whipple says, *op. cit.*, p. 83: "At the same time we found the subject had overlooked this element of his field of sensation, although his attention was directed with a certain strength toward it, so we marked his confidence as 0. This happened in cases where the judgments were so much effected by the difference of pressure as to be correct three times out of five." Griffing, *op. cit.*, p. 46, found the average of all correct guesses was 59 per cent. "In the case of some observers whose confidence was small, the percentage ran as high as 65 per cent. and 70 per cent."

[3] Professor Cattell's Psychophysical Contributions, *Arch. of Psychol.*, 1914, **30**, 60–74.

individual. One individual insisting on a rather high degree of clearness or confidence for his 'greater' or 'less' judgments will throw a large portion of the cases into the middle class, while another individual will follow slighter indications and throw similar cases into the plus or minus column. The individuals differ in their standards of confidence, but not necessarily in their power to perceive a difference. Now since it has been proved that the division of the middle class between the other two classes results in a good balance of correctness, the individual who takes the more chances or has the lower standard of clearness will get the better record." Woodworth therefore continues to advocate forcing the subject into the false position of inhibiting equal and doubtful judgments in order to make results of all subjects comparable. This end can probably be attained better by the method outlined in this paper. In regard to the equal judgments, as was said before, none of the subjects seem to have had difficulty in getting the positive perception of equality (except the one subject with the preëstablished set, and even he made many such equal judgments). The doubtful judgments are allowed, but not included in the results.[1] The results showed no indication of avoiding a positive decision or of lapsing into a state of mild 'aboulia,' as Warner Brown intimated would happen under such conditions.[2] If the doubtful judgments were allowed to remain with the equal judgments in the results, Woodworth would be justified in saying that "the boundaries of the middle class are fixed by each subject without objective control and almost invariably differ from subject to subject, etc.,"[3] for we saw how many categories the *d* judgments included. By counting only the very definite category of 'equality' in the sense of a decided impression of equality the boundary is very firmly fixed. By allowing the subjects who so desire to make doubtful

[1] G. W. Whipple repeated the doubtful tests until the observer made satisfactory judgments, but the number and distribution of these cases were recorded. *Op. cit.*, p. 417.

[2] The Judgment of Differences, *Univ. of California Publications in Psychology*, 1910, 1, 32–33.

[3] *Op. cit.*, p. 65.

judgments we gain two points. We keep the boundary sharply defined in his consciousness and by noting such judgments are able to make and allow for a change of attitude. These doubtful judgments tend to grow less with practice and would probably disappear in the latter half of a series of one hundred judgments. In other words, the method of impressing upon the subject the difference between judgments based on the actual perception of equality and those called equal because there was no judgment one way or the other probably had the beneficial effect of preventing too many equality judgments, and the knowledge that guess judgments would not stand as such kept the subject keyed to the proper amount of attention and probably prevented him from shifting his attitude.

Summary

1. The measurement of sensitivity for finger span with a standard of 5 cm. for the two most reliable subjects was .45 mm., or about 1 per cent., when one hand was used. It was in most cases higher when the stimulus was presented to the two hands, whether successively or simultaneously, but for one of the two most reliable subjects it changed very little. His M. S. remained relatively low for all three forms of presentation.

2. There was no decided agreement between subjective confidence and reliability. The best subject did have a high degree of confidence, but the second best had a low degree.

3. Decidedly more right judgments had a higher degree of confidence than wrong judgments.

4. When the subjects were forced to make judgments which had the quality of guesses, there were many more right than wrong answers.

A NEW OLFACTOMETRIC TECHNIQUE AND SOME RESULTS

BY HERBERT WOODROW AND BENJAMIN KARPMAN

University of Minnesota

The technique here to be described consists mainly in a new odorimetric method, which enables one to control accurately and easily the physical intensity of odoriferous stimuli. Besides a new odorimetric method, however, we have developed, by combining this method with the measurement of adaptation times, a new olfactometric technique, which should be of general usefulness in the study of the sense of smell. We shall first describe the odorimetric method, and explain its theoretical basis, and then give some results on adaptation time, showing in particular how adaptation time varies with the physical intensity of the odor.

Odorimetric Method.[1]—The fundamental principle of the odorimetric method is simply that the number of molecules given off to the air by any substance increases with its temperature. The number of odorous molecules in a given volume of atmosphere is here regarded as the physical intensity of the odor. At a given temperature, the number of molecules given off by an odorous substance in a closed space depends upon what is known as the vapor tension of the substance. The vapor tension increases with the temperature. Determinations of the vapor tensions for a wide range of temperatures have already been made by chemists in the case of a large number of odorous substances; so that about all the experimenter has to do is to regulate the temperature of the odor with which he wishes to experiment.

Briefly stated, the procedure is as follows: Some of the

[1] We are pleased to acknowledge indebtedness to Professor Hunter, of the chemistry department of the University of Minnesota, and to Professor Gibson, of the University of the Phillipines, for important suggestions, both as to method and apparatus.

odorous substance is placed at the bottom of each of a pair of Emmerling absorption tubes filled with glass beads. A moderate current of air, supplied by a water air-pump, is passed through these tubes and on into the nose of the subject. The temperature of the tubes containing the odorous substance is regulated by means of an Ostwald thermostat and accessories. In order to avoid the danger of precipitation during the transit of the odorous vapor from the Emmerling tubes to the nose, and also in order to avoid the unpleasantness which most odors produce when they become very strong, temperatures above 33° C. are not used. No other air is allowed to enter the subject's nose than that coming through the tubes. The subject breathes gently and regularly through his mouth, which is kept wide open. The air, in passing through the Emmerling tubes, becomes as saturated with the odorous vapor as the temperature permits. Just after leaving the Emmerling tubes, and just before entering the nose, the air is warmed, by being passed through a hot water bath, to a temperature of 33° C.; so that no matter what the temperature of the odorous substance in the Emmerling tubes, the mixture of air and odorous vapor entering the nose always has the same temperature.

Now, according to Avogadro's rule, equal volumes of gas or gas mixture, at the same temperature and pressure, have equal numbers of molecules. In all cases, then, that is, no matter what the substance used nor its temperature in the Emmerling tubes, the same number of molecules of gas mixture enter the nose in a given unit of time; for in all cases the air pressure and the temperature at the nose are constant. The percentage of the total number of molecules entering the nose composed by the odorous ones will be directly proportional to the vapor tension of the odorous substance in the tubes, no matter what the substance. The rise in temperature of the gases just before they enter the nose does not change the *proportion* of odorous molecules. Since, then, equal numbers of molecules of the gas mixture (air and odorous vapor) are entering the nose in each unit of time, and since the proportion of odorous molecules in this mixture is

directly proportional to the vapor tension of the odorous substance in the tubes, the absolute number of odorous molecules entering the nose in unit time is directly proportional to the vapor tension. Consequently, we may use as our units of intensity the vapor tension (expressed in millimeters of mercury). When the vapor tension doubles, the intensity doubles, that is, twice the number of odorous molecules enter the nose in given time. Or, from two different odorous substances, we will have a given unit of time the same number of molecules entering the nose, providing we keep these substances at such temperatures that their vapor tension is the same.

Since the phenomenon of vapor tension forms the basis of the preceding method, the theory of the method can be made clear only by a brief exposition of the concept of vapor tension. According to the kinetic theory of gases, the molecules of a gas are practically independent of each other and are briskly moving in all directions in straight lines. Because of the freedom of its molecules, a gaseous substance will expand, and fill any space presented to it. As it fills the vessel, the impact of the gas molecules on the walls of the vessel produces a certain pressure, which is the gaseous pressure. In the case of a liquid, the molecules still possess sufficient independence to slide over one another. In spite of the clinging together of the molecules of the liquid, some of them have sufficient motion to free themselves from their neighbors, and, leaving the liquid surface, become gaseous. The same is true, though to a lesser degree, of solids. Now, since the velocity of the molecules depends directly on the temperature, it is evident that as the temperature rises the number leaving the liquid will rise. These molecules which have left the liquid give rise to a pressure, which is called the vapor pressure. Hence the vapor pressure is a measure of the number of molecules of the liquid present above it, in a confined space.

If an odorous liquid be placed at the bottom of an enclosed vessel, the pressure of the molecules given off by it will then be exerted in all directions, against the liquid as well as in

other directions. After a certain time, owing to the vapor pressure, as many molecules will press back into the liquid as succeed in freeing themselves from it. The liquid and the vapor are then in equilibrium. Consequently, it may be said that the vapor tension of the liquid equals the vapor pressure.

Now, the number of molecules of a gas in a gas mixture is proportional to its partial pressure. In the gas mixture consisting of air and an odorous vapor, after equilibrium has been reached between the odorous vapor and its liquid, the partial pressure of the odorous vapor is equal to the vapor pressure, which, in turn, is equal to the vapor tension of the liquid at the given temperature. Hence the number of molecules of the odorous vapor in a mixture of air and vapor, in a confined space above the odorous substance, is proportional to the vapor tension of the odorous substance.

If two odorous vapors have the same partial pressure at the same temperature, then their molecules will be present in the mixture with air in equal numbers and in the same proportion or molecular concentration. (The same weights of substance will not be present, unless their molecular weights are the same.) If two odorous vapors have the same partial pressure, but at different temperatures, then they will not have equal numbers of molecules. They will, however, have the same molecular concentration, that is, the same proportion of the total number of molecules of the mixture of odorous vapor and air; and they may be made to have the same absolute number of molecules by equating the temperatures of the gas mixtures. In a gas mixture, the molecular concentration of the components remains the same with rise in temperature. If the temperature of a mixture be lowered, liquid will settle out, the amount depending, for a given change, on the shape of the vapor tension curve (vapor tension plotted against temperature). It is necessary, therefore, for our purpose, to equate the temperatures of our mixtures of air and odorous vapor by *raising* the temperature, at least that of the mixture with the lowest temperature. Otherwise, we would incur an error due to precipitation.

Actually, as already stated, we *raise* the temperature of all mixtures to 33° C., before they enter the nose.

Lastly, it should be kept in mind that the molecular weight of a substance has no effect on the number of molecules which are necessary to produce a given pressure. There are the same number of hydrogen molecules present in a unit volume of pure gas at a given temperature and pressure as there would be of any gas, however heavy. This is simply a restatement of Avogadro's rule.

The apparatus consists of well-known, inexpensive chemical pieces. It is made up, first, of a mechanism for injecting a steady current of pure air through the Emmerling absorption tubes; second, of the Emmerling tubes, which contain the odorous substance; third, of mechanism for controlling the temperature of the tubes; and fourth, of two glass tubes which conduct the odorous air from the Emmerling tubes through a hot water bath into the nose. The tubing used consisted exclusively of glass. The necessary junctions were effected by especially devised aluminum joints. These were made of two short aluminum tubes, each of which fitted over one of the two glass tubes it was desired to hold together. The two aluminum tubes were threaded so that one screwed into the other. Disconnecting the tubes was therefore almost as easy as unscrewing the tip of a fountain pen. The short aluminum tubes were attached, at the end away from the joint, to their glass tubes either by means of a paste of lead oxide or by means of adhesive tape, thus avoiding the use of rubber.

The air current was supplied by means of a water air-pump or waterblast connected with the laboratory hydrant. Before entering the Emmerling tubes, the air passed through two Novy jars. The first, containing a solution of potassium permanganate, served to rid the air of impurities and the second, containing calcium chloride, to rid it of moisture. Upon smelling the air as it came from the second Novy jar, no sign of an odor was perceptible.

The air next passed through the Emmerling absorption tubes, where it became saturated with the odorous vapor.

We began by using three Emmerling tubes, the first placed outside of the thermostat and the last two inside it. By preliminary trials, however, with propyl alcohol, it was found that the adaptation time was, on the average, very nearly the same whether we used two tubes or three. We consequently concluded that the use of two tubes was sufficient to bring about saturation, and thereupon discarded the tube outside of the thermostat.

The Emmerling tubes were placed in a large Ostwald thermostat, with motor-driven fan. The tank of the thermostat was filled either with ice and water or water only. The temperature was regulated automatically by means of a Reichert-Novy gas regulator and a Bunsen burner. One could probably get along with an ordinary tub and a supply of ice and hot and cold water.

After leaving the second Emmerling tube, the odorous air current passed immediately into a tube about one yard long but bent so as to fit into a pan of water about ten inches long. The water in this pan was kept at 33° C., by means of a Bunsen burner, regulated, when necessary, by hand. In the work with propyl alcohol, in which the thermostat temperatures used were all below room temperature, this hot water bath was omitted. Lastly, the air passed into a short tube, almost vertical, which entered the subject's nostril. The odorous air issued from this tube into the nose at a rate of 21.8 c.c. per sec. The internal diameter of the tube was 5.0 mm., and its thickness 1.0 mm. The monorhinic, or one-nostril, method, with the other nostril held closed by the finger, was preferred to the dirhinic, or two-nostril, because with it, the stream of air has a greater force, thus reducing the possibility of interference from the breathing (through the mouth) and assuring a continuous stimulation, during expiration as well as inspiration. Of course a stronger air current than the one we used could be employed, but it might have to be passed through a larger number of Emmerling tubes to assure complete saturation. The odorous particles blown into the nose probably are not carried by the air current directly into contact with the *regio olfactoria*. They are

carried close to it and travel the rest of the distance, an exceedingly short one, by diffusion. The diffusion rates of different odorous vapors, while not the same, would not be appreciably different for such a short distance.

The accompanying diagram shows the various parts of the apparatus and the way they were connected together. In this diagram, all parts of the thermostat, except the immersion tank, have been omitted.

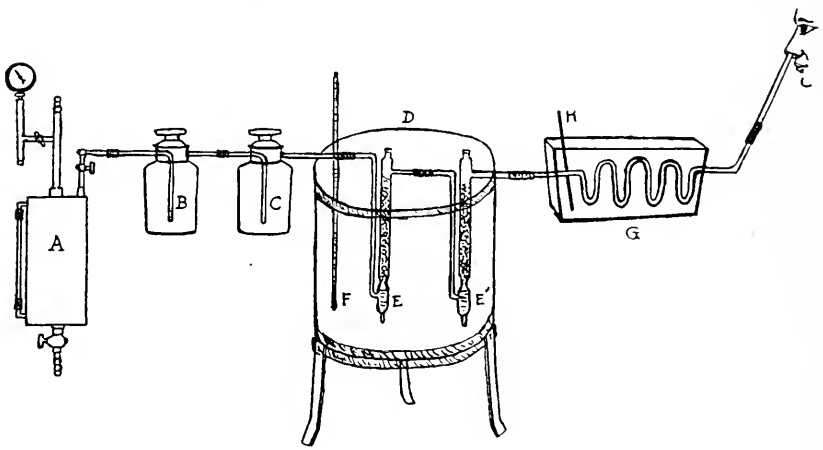


Fig. 1. Diagram of Apparatus.

A, air-blast.
B, Novy-jar, partially filled with solution of potassium permanganate.
C, Novy-jar, partially filled with calcium chloride.
D, thermostat tank.
E, *E′*, Emmerling absorption tubes filled with glass beads. (The odorous substance is placed at the bottom of these tubes.)
F, thermometer.
G, hot water bath.
H, thermometer.

The preceding odorimetric method is the only method so far described which enables one to stimulate the nose continuously for any desired time by means of an air current containing a known proportion of odorous molecules. It measures physical intensity of odor by means of units which are the same for different odorous substances, and this intensity does not vary with the duration of the stimulation.

Without going into a detailed criticism of the numerous

odorimetric methods employed in the past, we should like to point out one source of error, in addition to those commonly noted,[1] that appears to have been overlooked. This consists in an ignoring of the facts of vapor tension. For example, in the case of solutions, even if the solvent is odorless, the intensity of the odorous vapor at a given temperature does not vary, as often assumed, directly with the concentration of the solution. The vapor tension (or partial pressure) of ammonia in solution of water, at 19.9° C., is, for a 10.15 per cent. solution by weight, 80.6 mm. of Hg, while for a 23.37 per cent. solution, it is 302.4 mm. of Hg.[2] Thus, while the concentration of the solution increases 130 per cent., the physical intensity of the odorous vapor increases 275 per cent. In other cases, the increase in the intensity of the odor is less than proportional to the increase in the concentration of the solution.

Complications arise also with solids, as used, for example, in Zwaardemaker's method, a method which, in spite of the fact that it is known to suffer from serious errors,[3] appears to enjoy some popularity. It is based on the assumption that the mass of odorous particles given off from a solid body or the surface of a liquid, is, *other things equal*, proportional to the time and to the area of the surface exposed. This is true of a given substance in an unclosed or very large space. But in the application of this principle to an odorimetric method, Zwaardemaker makes use of an air current produced by inhalation through a tube composed of the odorous substance. This air current passes over the sections of the exposed surface in succession, and, as it does so, presumably becomes more saturated. At a given temperature, however, its saturation is limited. If the air, in passing over the first part of the surface, becomes saturated, its passing over

[1] For a history of previous odorimetric methods, see Gamble, *Amer. J. of Psychol.*, 1899, p. 98 ff. Among the common sources of error we may note the following: Adhesion of odorous particles to the walls of tubes or vessels; the use of odorous solvents; variation with the mode of breathing, and, in the case of bottled gases, with the time the bottle has been unstoppered; lack of control of the factor of diffusion (in Frölich's method); and incomparability of the intensities of different odors.

[2] Landolt and Börnstein, *Physicalische Chemische Tabellen*, 4th ed., 1912, p. 429.

[3] See Gamble *op,. cit.*, pp. 101–104.

additional surface will result in no increase in the intensity of the odor. Even though, as is probable, the air current in passing over the first part of the surface does not become saturated, its saturation in passing over a second portion of equal extent need not double, nor in passing over a third and fourth portion, triple and quadruple. Thus Zwaardemaker's method, in common with other methods, is theoretically unsound, because of neglect of the facts of vapor tension. The theory of vapor tension is fundamental to the understanding of the genesis of odors and must be taken into careful consideration by any odorimetric method that derives the odors from solids or liquids.

Adaptation Time and Intensity.—Having devised a satisfactory method for the control of the physical intensity of odor, we proceeded to study how the time required for adaption varies with the intensity. The first substance tried was propyl alcohol. With this odor four subjects were used. Later, experiments were made with two more substances, camphor and naphthalene, using three subjects. None of the subjects were smokers.

The vapor tensions of these substances, at the temperature used, is given in the adjoining table. Under the head of intensity, we have taken as units different vapor tensions in different cases. With propyl alcohol, we began with a temperature of 10° C., which gives a vapor tension of 7.4 mm. We arbitrarily called this two units of intensity, and then arranged our other temperatures so as to give a range of intensities from one to four units. With naphthalene and camphor we took as our unit of intensity a vapor tension of .05 mm. and adjusted our temperatures to give a range of from one to five units with naphthalene and from two to six with camphor.[1] The determinations of adaptation time were made in such order that the various intensities were mixed irregularly.

By way of practice, preliminary trials were made with

[1] The vapor tensions were taken from Landolt and Börnstein's *Physicalische Chemische Tabellen*, 4th ed., 1912. Inasmuch as the temperature is usually given in these tables by steps of five degrees (0, 5, 10, etc.), it was often necessary to interpolate in order to obtain the temperature required for the desired vapor tension.

each subject with each odor. The number of these preliminary trials was particularly large in the case of propyl alcohol, the first odor used. Propyl alcohol differed from the other odors in that with it there occurred a progressive shortening in the adaptation time with practice. With propyl alcohol, all trials were regarded as preliminary until this progressive shortening seemed to have ceased. In some cases no shortening was apparent after the second trial. In other cases it continued through five trials. With the other odors, where shortening with practice was not evident, each subject was given one preliminary trial with each intensity.

TABLE I.

SHOWING VARIATION OF VAPOR TENSIONS WITH TEMPERATURE

Substance	Temperature	Vapor Tension	Intensity
Propyl Alcohol	0.6° C.	3.7 mm. Hg.	1*x*
	10.0 "	7.4 " "	2*x*
	15.8 "	11.1 " "	3*x*
	19.7 "	14.8 " "	4*x*
Napthalene	11.0 "	.05 " "	1*x*
	24.5 "	.10 " "	2*x*
	31.0 "	.15 " "	3*x*
	34.4 "	.20 " "	4*x*
	36.9 "	.25 " "	5*x*
Camphor	10.0 "	.10 " "	2*x*
	20.0 "	.15 " "	3*x*
	25.8 "	.20 " "	4*x*
	29.5 "	.25 " "	5*x*
	31.5 "	.30 " "	6*x*

The preliminary results with propyl alcohol show that in some cases a certain amount of practice may be necessary before the subjects become reliable. They may have to learn to neglect many irrelevant sensations: taste, touch, burning, etc. With propyl alcohol, which, particularly at the higher intensities, produces a burning sensation, subjects gave the introspection that there was some difficulty in distinguishing between smell and touch and smell and taste. As the work continued, it became easier to discriminate the smell from the other sensations.

All the subjects reported that they could smell the odor

continuously, that is, during expiration as well as inspiration. This report seems natural enough, as the air current carrying the odorous vapor entered the nose at a steady rate, 21.8 c.c. per sec., while the subject breathed gently and exclusively through his wide-open mouth. One of the subjects, however, found that as the odor weakened with the progress of adaptation, it became much harder to be sure of it during expiration than during inspiration. All the subjects found it difficult

TABLE II.

ADAPTATION TIMES WITH PROPYL ALCOHOL

Intensity		1 *x*	2 *x*	3 *x*	4 *x*
Temperature		0.6°C	10.0°C.	15.8°C.	19.7°C.
Vapor Tension		3.7 mm.	7.4 mm.	11.1 mm.	14.8 mm.
Subj. Kn	Trial 1	102	144	180	240
" "	" 2	64	114	126	210
" "	" 3	110	96	115	240
" "	" 4	72	130	175	210
" "	" 5	55	144	156	120
" "	" 6	82	74	135	195
" "	Average	**81**	**117**	**148**	**203**
" "	*M. V.*	*17*	*23*	*22*	*29*
" Dy	Trial 1	24	126	135	255
" "	" 2	72	75	180	375
" "	" 3	78	126	130	105
" "	" 4	65	85	135	165
" "	" 5	51	96	150	210
" "	" 6	54	105	138	225
" "	Average	**57**	**102**	**145**	**223**
" "	*M. V.*	*15*	*17*	*13*	*62*
" Dl	Trial 1	115	85	155	285
" "	" 2	60	115	162	245
" "	" 3	45	90	192	105
" "	" 4	85	85	184	190
" "	" 5	186	130	156	145
" "	" 6	75	163	141	240
" "	Average	**94**	**111**	**165**	**201**
" "	*M. V.*	*37*	*28*	*15*	*56*
" Pn	Trial 1	72	95	175	270
" "	" 2	60	125	120	160
" "	" 3	90	85	100	180
" "	" 4	55	95	165	180
" "	" 5	70	95	150	135
" "	" 6	52	125	105	165
" "	Average	**67**	**103**	**136**	**182**
" "	*M. V.*	*10*	*15*	*27*	*29*
All Subjects	Average	**75**	**108**	**148**	**202**
" "	*Ave. M. V.*	*19*	*21*	*19*	*44*

to tell exactly when the odor was no longer present. They were instructed to say 'Now,' just as soon as they could decide for the first time that they were getting no sensation of odor. These instructions were due to the fact that in some of the preliminary trials we obtained statements as follows: "Twice it seemed as if the smell disappeared, only the burning sensation remaining. Then the smell appeared a third time before vanishing for good."

Table III.

Adaptation Times with Napthalene

Intensity		1x	2x	3x	4x	5x
Temperature		11.0° C.	24.5° C.	31.0° C.	34.4° C.	36.9° C.
Vapor Tension		.05 mm.	.10 mm.	.15 mm.	.20 mm.	.25 mm.
Subj. Ud	Trial 1	67	87	97	104	147
" "	" 2	73	84	94	98	129
" "	" 3	80	61	102	113	126
" "	" 4	65	79	86	98	131
" "	" 5	71	82	96	96	128
" "	" 6	76	77	98	114	136
" "	Average	**72**	**78**	**96**	**104**	**133**
" "	*M. V.*	*4*	*6*	*4*	*7*	*7*
" Fe	Trial 1	66	79	77	116	111
" "	" 2	84	74	88	94	131
" "	" 3	71	73	103	99	119
" "	" 4	72	69	96	103	129
" "	" 5	58	76	84	98	109
" "	" 6	53	73	89	104	124
" "	Average	**67**	**74**	**90**	**102**	**121**
" "	*M. V.*	*8*	*3*	*7*	*5*	*8*
" Kn	Trial 1	45	68	76	79	108
" "	" 2	50	72	86	93	103
" "	" 3	60	53	71	84	102
" "	" 4	49	67	73	69	101
" "	" 5	52	62	64	81	86
" "	" 6	48	62	74	77	100
" "	Average	**51**	**64**	**74**	**81**	**100**
" "	*M. V.*	*4*	*5*	*4*	*6*	*5*
All Subjects	Average	**63**	**72**	**87**	**96**	**118**
" "	*Ave. M. V.*	*5*	*5*	*5*	*6*	*7*

In the accompanying tables, Tables II.–IV., we give the adaptation time obtained at each trial. Six trials were made with each subject at each of the intensities. Each of the tables shows, for one of the three odors used, how adaptation

time varies with physical intensity. The absolute measure of the physical intensity in each case is the vapor tension, which is given at the heads of the columns in terms of millimeters of mercury. These vapor tensions may be taken as strictly comparable absolute measures of intensity, no matter what the odor. Figs. 2 and 3 show the variation in the average adaptation time with the physical intensity of the odors.

Table IV.

Adaptation Times with Camphor

Intensity		2*x*.	3*x*.	4*x*.	5*x*.	6*x*.
Temperature		10.0° C.	20.0° C.	25.8° C.	29.5° C.	31.5° C.
Vapor Tension		.10 mm.	.15 mm.	.20 mm.	.25 mm.	.30 mm.
Subj. Ud....	Trial 1........	102	160	184	231	265
" "	" 2........	97	133	175	216	252
" "	" 3........	98	142	186	193	236
" "	" 4........	104	152	169	208	228
" "	" 5........	84	134	142	222	202
" "	" 6........	82	132	145	230	215
" "	Average.......	**95**	**142**	**167**	**217**	**233**
" "	*M. V.*........	*8*	9	*15*	*11*	*18*
" Fe.....	Trial 1........	72	127	144	179	204
" "	" 2........	67	106	138	166	200
" "	" 3........	71	84	126	158	184
" "	" 4........	65	91	118	154	192
" "	" 5........	69	108	136	166	179
" "	" 6........	65	96	134	157	187
" "	Average.......	**68**	**102**	**133**	**163**	**191**
" "	*M. V.*........	*3*	*12*	*7*	*7*	*8*
" Kn....	Trial 1........	63	101	124	151	198
" "	" 2........	60	97	112	140	189
" "	" 3........	58	106	115	158	204
" "	" 4........	59	85	121	134	180
" "	" 5........	56	96	109	140	200
" "	" 6........	58	97	115	135	196
" "	Average.......	**59**	**97**	**116**	**143**	**193**
" "	*M. V.*........	*2*	*4*	*4*	*8*	*7*
All Subjects .	Average.......	**74**	**114**	**139**	**174**	**206**
" " .	*Ave M. V.*......	*4*	*8*	9	9	*11*

The data lead to definite conclusions. They show a marked increase in adaptation time with increase in the physical intensity of the stimulus. It is true that the variation in the adaptation time of any individual for any particular

intensity of odor is rather large. It is not so large, however, but that averages based on six determinations give fairly regular curves. This is seen by the fact that if the curves

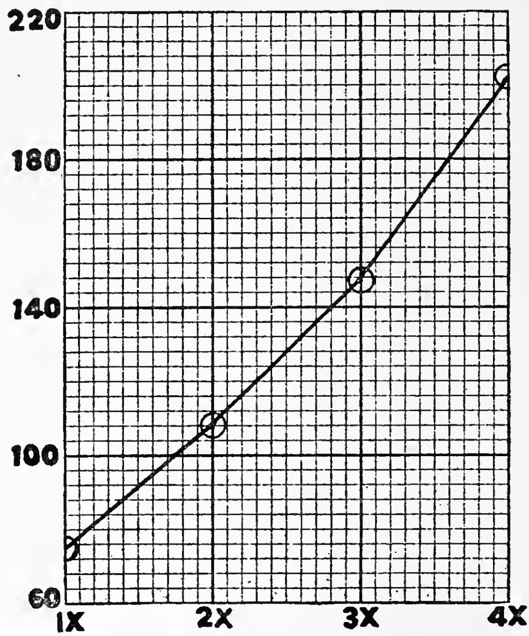


FIG. 2. Variation in Adaptation Time with Physical Intensity, Propyl Alcohol.

are plotted for the individual subjects, they will be found to have the same general slope. This fact justifies our averaging the results obtained with all subjects and plotting the curves of the averages, in order to determine the general relationship between adaptation time and physical intensity.

A study of these curves (Figs. 2 and 3) shows that they are fairly regular. They are at least regular enough to make it certain that adaptation time is a function of physical intensity and that this function can be determined with a high degree of precision. The data here presented are probably not sufficient in number nor in range to justify a final decision as to the mathematical law which prevails. It will be seen, however, that all three of the curves approximate straight lines, the formula for which is very simple. It is

that $t = K + k \cdot I$, in which t represents the adaptation time in seconds, K and k determinable constants and I the physical intensity. We may say, then, that adaptation time varies

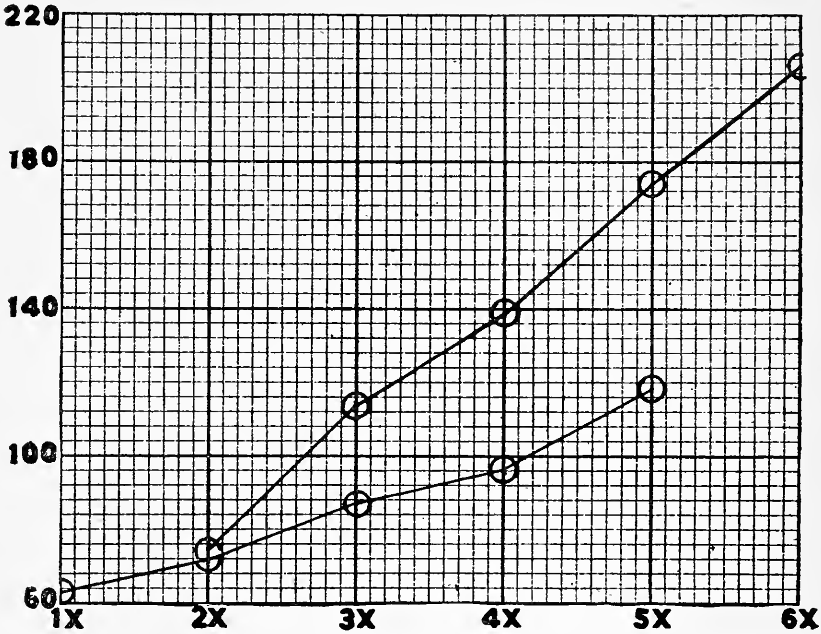


FIG. 3. Variation in Adaptation Time with Physical Intensity. Upper Curve, Camphor; Lower Curve, Naphthalene.

as a constant plus the product of another constant into the intensity of the stimulus; or, more simply, that *increasing the physical intensity of an odor by equal steps causes an increase in adaptation time by equal steps.* The *increase* in the one is directly proportional to the *increase* in the other. The very nature of this relationship suggests that it holds only within limits, at least that there is a lower limit, probably somewhere near the sensation threshold.

The following table shows the deviations of the adaptation times obtained experimentally from the most probable true ones on the assumption that $t = K + k \cdot I$. When the large mean variation and the small number of the individual determinations is kept in mind, it will be seen that the deviations are not serious.

TABLE V.

A COMPARISON OF OBTAINED ADAPTATION TIMES WITH THOSE CALCULATED BY THE LAW $t = K + k \cdot I$

Odor	Intensity	Obtained t	Calculated t	d
Propyl Alcohol...........	3.7 (1 x)	75	70	+ 5
$K = 28$	7.4 (2 x)	108	112	− 4
$k = 42$	11.1 (3 x)	148	154	− 6
	14.8 (4 x)	202	196	+ 6
Naphthalene..............	.05(1 x)	63	61	+ 2
	.10(2 x)	72	74	− 2
$K = 48$	.15(3 x)	87	87	0
$k = 13$	.20(4 x)	96	100	− 4
	.25(5 x)	118	113	+ 5
Camphor..................	.10(2 x)	74	76.5	− 2.5
	.15(3 x)	114	109.0	+ 5.0
$K = 11.5$	.20(4 x)	139	141.5	− 2.5
$k = 32.5$	.25(5 x)	174	174.0	0
	.30(6 x)	206	206.5	− 0.5

Intensity of stimulus has hitherto been regarded largely as though its sole mental function were intensity of sensation. We have shown that the duration of the sensation is also a function of intensity of stimulus, one which, on account of its easy measurement, as adaptation time, should be of fully as much interest as intensity of sensation in the study of the senses of smell, touch and temperature.

In conclusion, we desire to mention a few of the problems immediately suggested as suitable for study by the technique described in the preceding pages. One of these is the relation of adaptation time to intensity of sensation. Do two odors which have the same adaptation time have also the same intensity of sensation? We have not determined the answer to this question, but, on the basis of what observations we have made, we would incline towards an affirmative. At any rate, adaptation times should serve as valuable indirect measurements of sensational intensity.[1] Another interesting

[1] The relationship between the intensity of stimulus and the intensity of sensation, in the case of smell, has been studied by Gamble (The Applicability of Weber's Law to Smell, *Amer. J. of Psychol.*, 1899, 82–143). The interrelationship of the three factors, physical intensity, sensation intensity and adaptation time should be worked out not only for the sense of smell, but also for the temperature senses and the sense of touch.

question is the relationship of adaptation time to the classification of odors. Could we not accurately rank the qualitative resemblance of any number of odors to a standard by ranking them in the order in which their adaptation time is decreased by adaptation to the standard? For such work, it might be desirable first to equate the adaptation times of the different substances. This could be done by the present technique, aided by calculation. Using this technique, then, it might very well be possible to explore the odor-prism of Henning[1] and determine the correctness of its various dimensions. These problems are sufficient to serve as illustrations of the value of a sound odorimetric technique and of the usefulness which such a technique may acquire when combined with the measurement of adaptation times.

[1] Henning, Der Geruch, *Zsch. f. Psychol. u. Physiol. d. Sinnes.*, 1915, p. 254.

THE MEMORY VALUE OF MIXED SIZES OF ADVERTISEMENTS

BY HENRY F. ADAMS

University of Michigan

When advertisements are used in memory experiments, difficulties are encountered which are easily avoided by the use of more homogeneous materials. In the first place, the purpose of the advertising test is somewhat different from that of the ordinary memory experiment, for with advertisements, the desired result is the recall either of the name of the commodity or a description sufficient to identify it, or the recollection of some of its attractive qualities. Should any of these be remembered, the advertisement receives full credit.

Secondly, advertisements differ from each other in every conceivable way both on the objective and the subjective sides. The greater number of experiments which have been performed have dealt with the objective factors of size and frequency, undoubtedly because they are easier to control. Since advertisements differ also from the subjective standpoint, it is necessary to take some account of it as well. This is usually done by having a sufficient number of advertisements of the same size or of the same number of presentations so that it seems possible to disregard other differences and assume that variations due to interest factors will be compensated for in the averages. That this assumption is not unwarranted is indicated by the close similarity in the results of different investigators.

It might appear from the practical uniformity of results that the effects of the objective factors had been sufficiently determined and that further experimentation in that direction would be superfluous. But new considerations are constantly appearing which suggest further experiments.

The purpose of this paper is to give the results of an

experiment in which different combinations of sizes of advertisements were given four presentations. In securing the material, four advertisements of the same commodity were selected, but not all four were of the same area. There might be three full pages and one half; or two full and two half; or one full, one half and two quarter pages of the same commodity. Necessarily the different advertisements of the same article were not identical in content, but exhibited a considerable range of variation. In all, each combination of size was represented four times so that the following conditions were met:

4 firms using 3 full and 1 half page,
4 firms using 3 full and 1 quarter page,
4 firms using 2 full and 2 half pages,
4 firms using 2 full and 2 quarter pages,
4 firms using 2 full, 1 half and 1 quarter page,
4 firms using 1 full, 2 half and 1 quarter page,
4 firms using 1 full, 1 half and 2 quarter pages,
4 firms using 1 full and 3 half pages,
4 firms using 1 full and 3 quarter pages,
4 firms using 3 half and 1 quarter page,
4 firms using 2 half and 2 quarter pages,
4 firms using 1 half and 3 quarter pages,
11 firms using 1 full page.

It will be seen that every possible combination of four presentations, using mixed sizes, was employed. These advertisements were pasted upon cardboard the size of the standard magazine page, no two advertisements of the same commodity appearing on the same or neighboring pages. This dummy was handed to the subject, who was asked to look it through in the same way that he would the advertising section of a magazine, except that he was to be sure that he had seen every advertisement. If one did not interest him, however, he was told to skip it and go on to the next. He was requested to spend about 20 minutes on the dummy, but might take more or less time if he desired. In this way, 35 subjects were tested.

Before giving the results, it is necessary to outline briefly the other experiments which have been performed to test the effect of size and frequency of insertion of advertisements upon their memory value. The reasons for this will appear shortly.

1. *The Effect of Size.*—This topic has been investigated by Scott,[1] Strong,[2] Starch,[3] and Adams.[4] While some of the tests were carried on by the recognition method, others by the recall, the form of the curve and the ratios were quite similar, though the absolute amount remembered was somewhat higher with the recognition test.

Another difference was the material used. Scott, Starch, and Strong in certain of his experiments, used magazines, which prohibited the selection of unfamiliar advertisements. The result was that certain of the full-page and a lesser number of half-page advertisements received undue credit because they were already well known to the reader. Starch devised an ingenious correction for familiarity and found that a full page should be allowed 63 per cent. of its obtained credit, a half page 73 per cent. and a quarter page 100 per cent., but it seems probable that he overcorrected, thus depriving the larger sizes of a certain amount of deserved credit.

A third difference was the time which intervened between presentations. The usual method has been to give the entire series in a period of less than an hour and to test immediately. Strong, however, presented 4 dummies at intervals of a few minutes, 1 day, 1 week and 1 month. In the first three cases, the test took place 4 weeks after the first set of advertisements had been seen, in the last, 16 weeks after the first dummy had been read through. If no elaborate corrections are made for probable differences in memory value owing to the irregular lapse of time, and this ought to be rectified in the ratios themselves, the following figures appear:

[1] Scott, W. D., 'The Psychology of Advertising,' 168–169.

[2] Strong, E. K., Jr., Psychol. Rev., Vol. 21, 137 ff. J. of Exper. Psy., Vol. 1, pp. 319–338.

[3] Starch, D., 'Advertising,' pp. 30 and 48.

[4] Adams, H. F., *J. of Phil., Psychol., etc.*, 1916, 13, 141–152. 'Advertising and Its Mental Laws,' pages 226–237.

Interval	Size of Advertisements		
	¼ Page	½ Page	1 Page
None	1.00	1.94	3.29
1 day	1.00	1.53	2.39
1 week	1.00	2.24	3.60
1 month	1.00	1.47	2.23

In the other experiments, all the advertisements were contained in one dummy, so the results are more nearly comparable with Strong's first set of ratios. The average of these other tests is given below:

Size of Advertisements		
¼ Page	½ Page	1 Page
1.00	1.87	3.16

These ratios are significantly similar to Strong's first set. However, if correction is made for familiarity in the manner suggested by Starch, these figures appear:

Size of Advertisements		
¼ Page	½ Page	1 Page
1.00	1.37	1.98

With such correction, the memory value of size varies approximately as the square root of the area; without correction, approximately as the 1.15 root of the area. The true figure probably lies somewhere between the two values. That both sets of figures should approximate so closely to simple root curves is striking. Strong's figures indicate that size is less and less effective as the repetitions occur further and further apart.

2. *A. The Effect of Frequency of Insertion. Duplicates Used.*—Less work has been done on frequency than on size. The general method employed has been to prepare a dummy, or series of them, in which quarter, half and full pages appear one, two, and four times. In part of the experiments the advertisements have been presented together; in others varying intervals have existed between the successive appearances. Strong investigated the latter condition in the experiment already mentioned under the topic of size. Working his

results through in ratios, as was done before, these figures are obtained:[1]

Interval	Number of Presentations		
	1	2	4
None	1.00	1.52	1.84
1 day	1.00	1.57	2.05
1 week	1.00	1.64	2.35
1 month	1.00	1.19	1.53

Adams, whose advertisements were all shown in one dummy, obtained these ratios for 1, 2 and 4 presentations: 1.00, 1.49, 2.57. If these are averaged with the corresponding figures of Strong's, the values immediately below are found:

Number of Presentations		
1	2	4
1.00	1.51	2.20

These ratios cannot properly be corrected for familiarity, since they were obtained by giving one insertion a credit of 1.00, no matter what its size might be, and putting the other frequencies in relative terms. They are between the corrected and uncorrected ratios for size and show that the effectiveness of frequency varies approximately as the 1.7 root of the number of insertions. From these figures alone it is impossible to determine which is more effective, size or repetition. However, from Strong's figures, quoted above, it appears that, on the average, size has greater effect than frequency of presentation. The same point is brought out by a consideration of results from certain mail-order campaigns.

With the exception of the one month interval, Strong's data indicate that the longer the time which elapses between successive insertions of duplicated advertisements, the more effective the second and fourth insertion will be. This is quite the opposite of the findings for size.

[1] In the tables adapted from those given by Strong, the unit of presentation or of size is given throughout the value of 1.00, for we were interested in obtaining the relative effectiveness of multiples of the units. The subject matter of this paper does not make desirable a discussion of the effect of the varying intervals upon the memory value of the unit.

2. *B. The Effect of Frequency of Insertion. Variations Used.*—It has been found that variation has a much greater effect upon memory than duplication. The experiment[1] yielded the following ratios:

Frequency of Presentation		
1	2	4
1.00	2.63	4.05

Since only 40 subjects were employed in this investigation, another similar experiment was performed with different materials in which 122 persons were tested. The second experiment gave these ratios:

Frequency of Presentation		
1	2	4
1.00	2.21	4.15

These two tests, performed on different groups of students, with different material and the experimental work done by different men, produced results which show the same general tendency. If a weighted average of the two is determined, in which the number of subjects used in each experiment is taken into account, the following ratios are obtained:

Number of Presentations		
1	2	4
1.00	2.33	4.13

From these experiments, it is justifiable to state that the effectiveness of variation varies almost, but slightly more than, directly with the number of presentations.

These are the more important experiments which have been performed upon the memory value of advertisements. Is it possible to take what they show, accept it provisionally, and deduce further laws? It will be interesting to try. In order to do so, however, certain definite assumptions must be made. These will now be considered.

It will be recalled in connection with the memory value of size that two sets of ratios were given, one uncorrected for familiarity, the other overcorrected, and the true value

[1] Adams, H. F., *J. of Phil., Psychol., &c.*, 1916, 13, 151–152.

was to be found between the two. Starch, who grants that his method was a makeshift, had his subjects write a list of all the commodities with which they were familiar, through advertising or any other means. The number of times each commodity was mentioned was then subtracted from the number of times it was remembered in the experiment and the remainder was supposed to be the true value of the advertisement. This surely is overcorrecting, for even advertisements which were wholly unfamiliar would have some memory value with these subjects and Starch's method overcorrects by at least that amount. Since Strong found that the average firm was remembered by about 50 per cent. of those experimented upon immediately after reading the dummy, it seems fair to split the difference between the results obtained with and without correction. This is done by giving a half-page advertisement 86.5 per cent. and a full page 81.5 per cent. of the obtained credits. When this is done, the following ratios are obtained for size:

Size of Advertisement		
¼ page	½ page	1 page
1.00	1.62	2.57

It must be remembered that the figures refer to those found when the presentations and the test all took place in a few minutes. These ratios vary approximately as the 1.45 root of the area. For the theoretical considerations which follow they will be taken as the working basis. There is justification for them because we are fairly sure, both from experimental evidence and from business returns, that doubling or quadrupling the size produces greater effect upon memory than a corresponding increase in the number of presentations with duplicated advertisements.

In the table immediately following is given a summary of the findings which will be used:

	Units of Stimulation			Efficiency
	1	2	4	
Size	1.00	1.62	2.57	1.45 root
Duplication	1.00	1.51	2.20	1.70 root
Variation	1.00	2.33	4.13	Almost directly

Since, in the experiment which serves as a basis for this paper, the advertisements of the same commodity were variations and also differed in size, it is possible to take the two tables dealing with these factors and put them together. It will then be found that certain very definite theoretical conclusions will appear. This table follows:

Size	Frequency of Presentation			
	1	2	3	4
¼	1.00	2.33	3.30	4.13
½	1.62	3.78	5.35	6.70
1	2.57	6.00	8.48	10.60

The figures are obtained by multiplying the ratios for size, in the perpendicular column under 1, by those for frequency, in the horizontal column opposite ¼. The values for frequency show the effect of 1, 2, 3 and 4 insertions of the advertisement. The effect of the second insertion may be determined by subtracting the ratio of one insertion from that of two; of the third, by subtracting that of two from that of three; and so on for the rest of the combinations. When this is done, these figures are obtained:

Size	Appearance Number			
	1	2	3	4
¼	1.00	1.33	0.97	0.83
½	1.62	2.16	1.57	1.35
1	2.57	3.43	2.48	2.12

If we wish to determine the probable memory value of a full page, followed by a half, followed by a quarter, we find in the table the effectiveness of the first presentation of the full page, 2.57, of the second presentation of the half page, 2.16, of the third presentation of the quarter page, 0.97, and add them together, getting 5.70. Similarly, if we wish to invert the order and determine the utility of a quarter, followed by a half, followed by a full page, we add the corresponding numbers in the table, 1.00, 2.16, 2.48, the sum being 5.64. This shows that the former order of presentation is slightly better than the latter.

But we have entirely disregarded one factor, and a very important one, too. It seems entirely plausible that a half page coming after a full page should be more effective than the same half page following a quarter page advertisement of the same commodity. This would be necessary if we take account of the conditions of attention, for we know in a general way that the direction of attention is determined by the relative vividness of the ideas, or by the relative permeability of brain tracts. A full page must, on the average, decrease the resistance of synapses more than a quarter page and leave the brain better prepared for attending a second time to the same or a similar object. Consequently, a half page coming after a larger size must have a higher value than one following a smaller size. The actual amount is conjectural in the present state of our knowledge. As a working basis, we have assumed that an advertisement preceded by one of double the area should receive an increase of 10 per cent. of its actual value, whereas, when the foregoing advertisement is half the area, it gets a corresponding reduction of 10 per cent. Likewise, when a quarter page followed a full, it received an arbitrary increase of 20 per cent. and a full page coming after a quarter, a decrease of 20 per cent. Similarly, when a full page was followed by 3 half pages, the first half page received an increase of 10 per cent., the second of 5 per cent. and the third of 2.5 per cent. These figures are intended to be merely illustrative. It is hoped that from a series of experiments now going on more definite and trustworthy values may be assigned. Some correction, however, is necessary.

If we do the examples given above, taking account of the corrections just outlined, the following is obtained:

Full page	2.570
Half page	2.16 + .216 = 2.376
Quarter page	0.97 + .097 = 1.067
Sum	6.013
Quarter page	1.000
Half page	2.16 − .216 = 1.944
Full page	2.48 − .248 = 2.232
Sum	5.176

This way of figuring out the results shows that the climax order, starting small and increasing the size, is decidedly less effective than the anticlimax order, an interesting difference between the art and the science of advertising.

It will be interesting to determine, on the basis of these figures, what we should expect as the result of our experiment, and then show the values actually obtained. In the experiment, however, no attention was paid to the order in which the different sizes of advertisements appeared, so, to make the two sets of figures at all comparable, it is necessary to find the maximum and minimum expectancy values and then observe whether the results obtained in the experiment lie between the two curves or a short distance to one side. In the following set of tables will be given the probable values of the different combinations of four presentations of the various sizes used in the experiment (pp. 458–9).

If these tables are based upon a correct theory, it is demonstrated that the order in which mixed sizes are presented will have a considerable effect upon the memory value of the advertisements of a firm. Since there is a close correlation between the memory value of mail order advertisements and their inquiry-producing power, it appears likewise that the actual effectiveness of a campaign will depend upon the order of presentation of the different sizes of advertisements which are used. From the tables it is possible to draw two quite definite conclusions.

1. The maximum efficiency is obtained when the largest advertisements appear first, followed by the smaller ones in descending order of size.

2. The minimum effectiveness is produced when the smallest advertisements are seen first, followed by the larger in ascending order of size.

These points ought to be of considerable value to the practical advertiser, likewise to the educator, for they demonstrate that the same amount of stimulation may produce vastly different results, one method showing a gain of from 10 per cent. to 20 per cent. in efficiency, the other a corresponding loss.

TABLE I

	1	2	3	4
3 full—1 half:				
FFFH	2.57	6.00	8.48	9.97
FFHF	2.57	6.00	7.73	9.64
FHFF	2.57	4.95	7.18	9.19
HFFF	1.62	4.71	7.06	9.13
3 full—1 quarter:				
FFFQ	2.57	6.00	8.48	9.48
FFQF	2.57	6.00	7.16	8.86
FQFF	2.57	4.17	6.15	8.06
QFFF	1.00	3.74	5.97	7.98
2 full—2 half:				
FFHH	2.57	6.00	7.73	9.15
FHFH	2.57	4.95	7.18	8.67
FHHF	2.57	4.95	6.60	8.51
HHFF	1.62	3.78	6.01	8.02
HFHF	1.62	4.71	6.44	8.35
HFFH	1.62	4.71	7.07	8.56
2 full—2 quarter:				
FFQQ	2.57	6.00	7.16	8.07
FQFQ	2.57	4.17	6.15	7.15
FQQF	2.57	4.17	5.24	6.94
QQFF	1.00	2.33	4.31	6.22
QFQF	1.00	3.74	4.90	6.60
QFFQ	1.00	3.74	5.97	6.97
2 full—1 half—1 quarter:				
FFHQ	2.57	6.00	7.73	8.64
FFQH	2.57	6.00	7.16	8.38
FHQF	2.57	4.95	6.02	7.72
FQHF	2.57	4.17	5.58	7.28
FHQF	2.57	4.95	7.18	8.18
FQFH	2.57	4.17	6.15	7.64
HFFQ	1.62	4.71	7.07	8.07
HFQF	1.62	4.71	5.87	7.57
HQFF	1.62	3.08	5.06	6.97
QHFF	1.00	2.94	5.17	7.08
QFHF	1.00	3.74	5.47	7.38
QFFH	1.00	3.74	5.97	7.46
1 full—2 half—1 quarter:				
FHHQ	2.57	4.95	6.60	7.51
FHQH	2.57	4.95	6.02	7.24
FQHH	2.57	4.17	5.58	6.86
HHFQ	1.62	3.78	6.01	7.01
HHQF	1.62	3.78	4.85	6.55
HFHQ	1.62	4.71	6.44	7.35
HQHF	1.62	3.08	4.49	6.40
HFQH	1.62	4.71	5.87	7.09
HQFH	1.62	3.08	5.06	6.55
QHHF	1.00	2.94	4.43	6.34
QHFH	1.00	2.94	5.17	6.66
QFHH	1.00	3.74	5.47	6.89

	1	2	3	4
1 full—1 half—2 quarter:				
FHQQ	2.57	4.95	6.02	6.89
FQHQ	2.57	4.17	5.58	6.49
FQQH	2.57	4.17	5.24	6.46
HFQQ	1.62	4.71	5.87	6.78
HQFQ	1.62	3.08	5.06	6.06
HQQF	1.62	3.08	4.15	5.85
QQHF	1.00	2.33	3.74	5.65
QQFH	1.00	2.33	4.31	5.80
QHFQ	1.00	2.94	5.17	6.17
QHQF	1.00	2.94	4.01	5.71
QFHQ	1.00	3.74	5.47	6.38
QFQH	1.00	3.74	4.90	6.12
1 full—3 half:				
FHHH	2.57	4.95	6.60	7.98
HFHH	1.62	4.71	6.44	7.86
HHFH	1.62	3.78	6.01	7.50
HHHF	1.62	3.78	5.35	7.26
1 full—3 quarter:				
FQQQ	2.57	4.17	5.24	6.11
QFQQ	1.00	3.74	4.90	5.81
QQFQ	1.00	2.33	4.31	5.31
QQQF	1.00	2.33	3.30	5.00
3 half—1 quarter:				
HHHQ	1.62	3.78	5.35	6.26
HHQH	1.62	3.78	4.85	6.07
HQHH	1.62	3.08	4.49	5.77
QHHH	1.00	2.94	4.43	5.74
2 half—2 quarter:				
HHQQ	1.62	3.78	4.85	5.72
HQHQ	1.62	3.08	4.49	5.40
HQQH	1.62	3.08	4.10	5.32
QQHH	1.00	2.33	3.74	5.02
QHQH	1.00	2.94	4.01	5.23
QHHQ	1.00	2.94	4.43	5.34
1 half—3 quarter:				
HQQQ	1.62	3.08	4.10	4.94
QHQQ	1.00	2.94	4.01	4.88
QQHQ	1.00	2.33	3.74	4.65
QQQH	1.00	2.33	3.30	4.52

We now return to a consideration of the experiment. It will be recalled that when it was started, there was no idea that the order of presentation would have any effect upon the outcome, so in the arrangement of the dummy, no attention was paid to this point. Consequently, in giving the results and showing the relation between the obtained and theoretical values, the maximum and minimum theoretical figures and the obtained ones will be shown, first in a table and then in a curve. The table follows:

Arrangement of Advertisements	Ratios of Av. Number of Mentions	A. D.[1]	Theoretical	
			Max.	Min.
3 full—1 half	9.55	2.50	9.97	9.13
3 full—1 quarter	8.57	8.00	9.48	7.98
2 full—2 half	7.81	8.75	9.15	8.02
2 full—2 quarter	7.16	5.00	8.07	6.97
2 full—1 half—1 quarter	8.41	1.75	8.64	7.07
1 full—2 half—1 quarter	5.03	5.25	7.51	6.34
1 full—1 half—2 quarter	5.52	6.50	6.89	5.65
1 full—3 half	7.57	3.00	7.98	7.26
1 full—3 quarter	6.26	2.50	6.11	5.00
3 half—1 quarter	6.51	6.00	6.26	5.74
2 half—2 quarter	5.03	4.00	5.72	5.34
1 half—3 quarter	4.12	1.75	4.94	4.52
1 full	2.12	1.60	2.57	2.57

In curve I. the same results are shown. When the values are plotted graphically, the very general similarity of the expected and the obtained results is clear. All but one of the points on the experimental curve lie close enough to the theoretical curve to be explained by the probable error of the experiment, which must be moderately high with only 35 subjects and 4 firms for each arrangement of advertisements. In the two curves, also, the directional change is the same in every place but one. While the results of the experiment do not absolutely prove the theoretical point, they render it extremely probable and open up an entirely new field for experimentation in the psychology of advertising.

An additional confirmation, while not completely conclusive, still points in the same direction. Shryer[2] gives, scattered through his book, a list of 9 insertions of advertisements followed by a second which was, judging from the

[1] It was pointed out by Dr. Froeberg that the A. D., in such an experiment, is practically useless. The A. D. may be worked out in either of two ways: First, by obtaining the average memory value for each advertisement and its A. D. By this method, the full page advertisements would have the smallest A. D., the half page next, and the quarter page advertisements the largest, for they are remembered by the fewest persons. Second, it is possible to group together all the firms advertising with 3 full and 1 half pages, in this case 4, obtaining the average and the A. D. This is the method which was employed, simply because some measure of variability is demanded in experimental work. But the A. D., to be of any value, presupposes material which is homogeneous, of equal difficulty. In these experiments, the advertisements are not of equal difficulty except by chance. Consequently, the A. D. is a product of two independent factors; the difficulty of the material, and the variability of the subjects.

[2] Shryer, W. A., 'Analytical Advertising.'

cost of space, 2.25 times the area of the first. Determining the probability in terms of his statistics for size and frequency, we find the expectancy value of the two insertions to be 2.575,

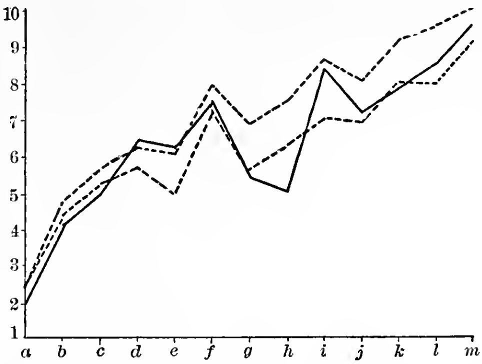


FIG. 1. Solid line: experimental results. Dotted line: maximum and minimum theoretical results.

a = 1 full,	g = 1 full—1 half—2 quarter,
b = 1 half—3 quarter,	h = 1 full—2 half—1 quarter,
c = 2 half—2 quarter,	i = 2 full—1 half—1 quarter,
d = 3 half—1 quarter,	j = 2 full—2 quarter,
e = 1 full—3 quarter,	k = 2 full—2 half,
f = 1 full—3 half,	l = 3 full—1 quarter,

m = 3 full—1 half.

and the actual value obtained from running the advertisements, 2.6000.

Still another bit of indirect confirmation of the theory is offered in curve II.

Fig. 2, *B*, shows the results obtained by using the values for size and frequency of insertion secured by certain mail order campaigns. Size was shown to have an efficiency of 1.00, 1.53 and 2.25 respectively for quarter, half and full page. Likewise, 1, 2, 3 and 4 insertions were discovered to give the following ratios: 1.00, 2.08, 2.78, 3.46. Using exactly the same method that was employed in the earlier parts of the paper, curve *B* was obtained. This curve is not guess work; it is based upon actual returns and a sufficient number of them to make it seem quite trustworthy. In giving results for frequency of insertion of advertisements in business, it is impossible to differentiate absolutely between duplication and variation, for both are probably effective. Fig. 2, *B*, then, would represent a mixture of the two.

Fig. 2, *A*, gives again the maximum theoretical results for variation, *C* for duplication worked out in terms of the tables on pages 458–9. Since curve *B* is a product of both of these factors, it ought to lie between *A* and *C*, and it does, almost equidistant from each all the way through, showing the same peculiarities, the same changes in direction, in fact, exhibiting the same tendencies throughout. The striking fact is the great degree of similarity in the forms of the various curves. The hypothetical ones, no matter whether for varia-

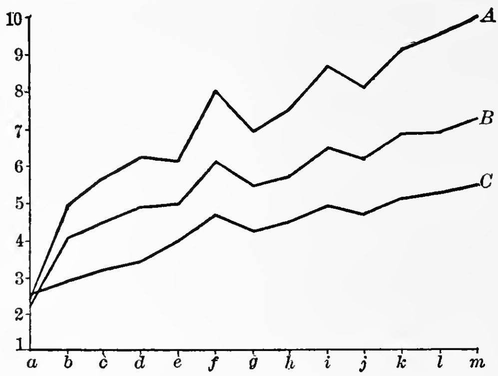


Fig. 2. The letters represent the same combinations as in Fig. 1.

tion or duplication, show the same general tendencies, though with differences of amount, as would be expected. The practical curve, *B*, is, to all intents and purposes, the mean between the two theoretical curves and the experimental curve is a very close approximation to *A*.

We may take it, then, that our main thesis is demonstrated in three ways, theoretically, practically and experimentally, viz., that using mixed sizes of advertisements may be made much more effective than using advertisements all of the same size. The second thesis seems almost equally probable, viz., that the order of distribution of the different sizes plays a large part in determining the efficiency of the total space employed.

Since there is such a close resemblance between the experimental results and the expectation curve, we shall use the latter as a basis for drawing further conclusions.

We do this because the expectation curve is based upon the results of a tremendous number of subjects as compared with the number used in our experiment.

If we take the combinations of sizes and show their relative effectiveness for different total areas according to the maximum theoretical curve, the following table is procured:

1 page	
4 quarters	4.13
1 half—2 quarters	4.10
1 full	2.57
1¼ pages	
1 half—3 quarter	4.94
2 half—1 quarter	4.85
1 full—1 quarter	4.17
1½ pages	
3 half	5.35
2 half—2 quarter	5.72
1 full—1 half	4.95
1 full—2 quarter	5.24
1¾ pages	
3 half—1 quarter	6.26
1 full—1 half—1 quarter	6.02
1 full—3 quarter	6.11
2 pages	
4 half	6.70
1 full—2 half	6.60
1 full—1 half—2 quarter	6.89
2 full	6.00
2¼ pages	
2 full—1 quarter	7.16
1 full—2 half—1 quarter	7.51
2½ pages	
2 full—2 quarter	8.07
2 full—1 half	7.73

There is no need to carry the table further, because all the other combinations are given in the preceding tables or curves. A brief inspection of the figures will show the most efficient combinations. The general conclusion is that when

variations are used, repetition is much more effective than size, and greater effectiveness will probably be obtained by using small, varied advertisements frequently inserted than by employing larger ones with less frequency. This is, of course, from the theoretical standpoint and takes no account of such considerations as the added cost of preparing new advertisements.

Where duplication is employed, provided anyone ever does use duplicated advertisements of different sizes, quite different results would be reached, for it will be remembered that with them, size is more effective than repetition. The exact amount of difference in any combination can be figured out by preparing tables like those given above.

In the practical situation, most advertising is a mixture of variation and duplication. From the standpoint of memory, what the effects of such a mixture will be is entirely problematical, but should be capable of experimental investigation with little difficulty. Certain practical results have already been given in Fig. 2, *B*, and it is probable that the memory curve would run closely parallel to it.

In conclusion, the results of psychological tests on advertisements, the results of the mail order campaigns which are accessible, these show that the effects of the objective factors follow rather simple root curves. The actual elevation differs somewhat. The results of certain other advertising campaigns, when plotted, follow definite power curves, the difference in the form of the curve resulting from differences in the kind of commodity.

These power and root curves are obtained by averaging or combining large numbers of records. It is not claimed that any one, or any dozen, insertions will lie along the curve, but in the long run, it will be approximated. This suggests a comparison with the tables which are used by life-insurance companies, which will not hold in any one case except by accident, but which, on the average, enable the life-insurance companies to make tremendous profits. It seems possible that advertising may in time become as scientific and as definite as insurance. But first, the advertisers must devise

some means of measuring the efficiency of their advertisements. Could they determine whether, for them, size and frequency followed definite power or root curves, they would have made a definite advance and could undoubtedly avoid considerable waste in conducting their campaigns. Even if they could not find root or power curves for the effectiveness of their advertisements, could they demonstrate the relative efficiency of mixed sizes, or mixed duplication and variation, they ought to be able to infer from their tables the type of curve which their advertising follows.

CHILDREN'S SENSE OF HARMONIES IN COLORS AND TONES

BY J. F. DASHIELL

Oberlin College

Object

Many students have tested children's ability to perceive colors,[1] but only a few have studied their affective preferences in single colors,[2] and fewer still their preferences in color combinations.[3] So far as the writer is aware, no systematic study has been made of children's preferences in musical tone combinations or intervals. A preliminary object of the present study was first to get the writer's own results on the subjects here observed as regards single color judgments to check up with other investigations. But the main interest was in learning various points about *children's capacity to judge in terms of color combinations or harmonies;* and lastly to treat in a similar way if possible *their judgments of tone intervals or harmonies.*

Materials and Methods

212 kindergarten children were used as the subjects, taken from nine Minneapolis kindergartens, both public and private, and including children of all classes. The experiments were given with the help of teachers-in-training. For purposes of comparison, the same observations were made upon 126 sophomores in the writer's classes in general psychology at the University of Minnesota.

[1] Baldwin, Binet, Bullough, Garbini, Lobsien, McDougall, Marsden, Monroe, Myers, Nagel, Preyer, Shinn, Wolfe, Ziehen.

[2] Engelsperger and Ziegler, 'Weitere Beiträge z. Kenntnis d. phys. u. psych. Natur d. sechsjährigen in d. Schule eintretenden Kindes,' *Experimentelle Pädagogik*, 1906, 2, 49–95. Holden and Bosse, *Archiv. of Ophthal.*, 1900, **29**, number 3. Winch, W. H., Color Preferences of School Children, *Brit. J. Psychol.*, 1909, 3, 42–65.

[3] Dobbie, W. J., Experiments with School Children on Color Combinations, *Univ. of Toronto Studies*, 1900, 1, 253–267. Wundt, W., Die Empfindung d. Lichts u. d. Farben, *Phil. Studien*, 1888, 4, 311–389.

In the investigation of single color preferences the six standard kindergarten colors (Milton Bradley Co.) were used —red, orange, yellow, green, blue, violet, corresponding, in Ridgway's classification, to colors 2i, 8–, 21–, 34–, 51–, 61h, respectively.[1] Rectangles of 13 mm. by 45 mm. were cut from paper of the six colors and pasted in the centers of neutral gray (Milton Bradley Co.) cardboards, 80 mm. by 90 mm. A modified order-of-merit method was adopted for the test, the child picking out the card bearing the color liked best from the whole number on the table, the experimenter removing it, the child picking the best of those remaining, etc. In the study of preferences in color combinations the same kind of cards was used, with same colors mounted in pairs, there being six pairs used—red-green, red-blue, orange-yellow, orange-green, yellow-violet, blue-violet. The order-of-merit method was again adapted to the situation.

In the experiments on tone harmony preferences, the experimenter, seated at the piano, sounded five musical intervals one at a time, asking before and after each, "Do you like this sound?" or "Do you think this a pretty sound?" Repetitions were allowed. The child's judgments in 'Yes' or 'No' were recorded for each interval. The intervals used were:

Major third,	c^1–e^1,	264–330 vibrations;	
" seventh,	c^1–b^1,	264–495	" ;
" fifth,	c^1–g^1,	264–396	" ;
Octave,	c^1–c^2,	264–528	" ;
Minor second,	c^1–$d^{\flat 1}$,	264–281.6	" .

Results

I

The Preliminary Test for Single Color Preferences.—In the tables is shown first the total number of times each color was given first choice, second choice, etc. Then, by multiplying the number of times a given color is placed in first

[1] R. Ridgway, "*Color Standards and Color Nomenclature.*" published by author, Washington, 1912.

rank by the figure 1, in second rank by the figure 2, etc., and adding the products, one obtains a figure representing inversely that color's value as resulting from all its placings. This is given as the total value (T. V.). On the basis of the total values, the six different colors can then be given their final order of preference or rank (R.).

TABLE I

Kindergarten Girls and Boys (212)								
	1	2	3	4	5	6	T. V.	R.
Red	45	34	29	31	37	36	725	2
Orange	28	36	37	35	40	36	767	6
Yellow	36	31	41	29	36	39	751	3
Green	29	33	38	42	37	33	760	5
Blue	40	42	34	42	28	26	690	1
Violet	34	36	33	33	34	42	759	4

The indefiniteness and small reliability of such measurements is at once evident, as in the difference between value 690 for blue (ranked first) and value 767 for orange (ranked sixth). One result of interest aside from the very slight differences given in the table is that the table for girls only and that for boys only shows agreement in giving first place to blue and fourth place to yellow. The difference of color given second place is striking.

TABLE II

Kindergarten Girls (107)			Kindergarten Boys (105)		
	T. V.	R.		T. V.	R.
Red	381	5	Red	344	2
Orange	416	6	Orange	351	3
Yellow	374	4	Yellow	377	4
Green	367	3	Green	393	5
Blue	348	1	Blue	342	1
Violet	361	2	Violet	398	6

The tables in a general way are in some agreement with the results of earlier studies.[1] The placing of red by the

[1] Winch for girls in (English) standards II. to VII. gives the following order of preference: blue, red, white, green, yellow, black; for boys: blue, red, yellow, green, white, black. Holden and Bosse find red preferred by infants to 3 years of age, with blue predominating more and more thereafter. Engelsperger and Ziegler find 6-year-olds giving decided preference to purple, dark blue, and violet.

girls is the only surprising fact. This may be due in part to the particular shade of red used. Later it will be observed that the college girls restore red to its expected place.

The correct interpretation of the close grouping of the total values for the respective colors would be important. This may be an indication that the order-of-merit method is not adequate for studying preferences of children. On the other hand, it more probably indicates that preferences in young childhood, however decided or undecided at the time, are comparatively irregular and probably inconstant, and are to be understood best in connection with the instability in all phases of childhood. It is further evidence for the lack of mental and nervous organization that is to be recognized as fundamental here.

When we compare the results above with the results of similar observations upon college sophomores interesting points come out.

TABLE III

Sophomore Girls and Boys (126)								
	1	2	3	4	5	6	T. V.	R.
Red	38	26	17	19	17	9	356	2
Orange	7	12	17	21	27	42	553	6
Yellow	9	16	12	31	31	27	508	5
Green	20	19	31	31	16	9	409	3
Blue	36	32	24	13	11	10	339	1
Violet	16	22	24	11	24	29	470	4

A very obvious difference is that here there is a more clear-cut order of preference, the total values of the six colors showing larger differences between them. Thus, between the 339 for blue (ranked first) and the 553 for orange (ranked sixth) we find a more substantial difference than between 690 and 767 above. Is this a matter of self-organization through experience, making one's maturer judgments proceed from a firmer physiological and psychological basis, or is it a matter of social control in the form of innumerable social agencies that tend to make individuals uniform? Put differently, the question is: is this greater uniformity indicative of a similarity of structure and innate tendency that shows itself with time, or is it indicative merely of the similarity of social forces that help determine development?

Aside from this difference of reliability, it is interesting to observe a fairly close correlation between the order of preference given by the sophomores and that by the child. With the exception of yellow and green, which are reversed, the orders are exactly the same: Blue, red, green or yellow, violet, yellow or green, orange. This might add a little strength to the first type of theory referred to.

Sex differences among the sophomores were indicated: the difference between the choice of red or blue for first and second place, and the difference between orange and violet for fourth and sixth.

TABLE IV

Sophomore Girls (92)			Sophomore Boys (34)		
	T. V.	R.		T. V.	R.
Red	255	1[1]	Red	101	2
Orange	419	6	Orange	134	4
Yellow	361	5	Yellow	147	5
Green	304	3	Green	105	3
Blue	271	2	Blue	68	1
Violet	321	4	Violet	149	6[2]

Comparing Tables IV. and II., the sex differences appear in general similar at the two ages, with the striking exception of the girls' low ranking of red in kindergarten and high ranking in college.

II

In the test for color combinations the kindergarten children showed decidedly little definiteness of choice.

TABLE V

Kindergarten Girls and Boys (210)								
	1	2	3	4	5	6	T. V.	R.
Red-green	47	29	35	34	33	32	703	1
Red-blue	38	35	30	44	37	26	715	2
Orange-yellow	42	35	35	26	31	41	722	3½
Orange-green	18	33	42	42	43	32	785	6
Yellow-violet	31	46	32	35	33	33	722	3½
Blue-violet	34	32	36	29	33	46	763	5

[1] Compare with Winch's observation that 41 women teachers ranked red low: blue, green, white, red, yellow, black.

[2] Compare Jastrow's report (Popular Æsthetics of Color, *Pop. Sci. Mo.*, 1897, **50**, 361–368) that adult men rank blue decidedly first with red second and that adult women rank red first with blue second.

Differences between boys and girls show in the placing of the first three cobminations.

TABLE VI

Kindergarten Girls (106)			Kindergarten Boys (104)		
	T. V.	R.		T. V.	R.
Red-green	368	4	Red-green	335	1
Red-blue	360	1	Red-blue	355	2
Orange-yellow	363	2	Orange-yellow	359	4
Orange-green	394	6	Orange-green	391	6
Yellow-violet	366	3	Yellow-violet	356	3
Blue-violet	375	5	Blue-violet	388	5

The extremely close grouping of the children's judgments of the color combinations is the striking fact of this set of observations. And as one casts his eye over the detailed results he is struck even more by the same thing. The children were grouped according to their nine different kindergartens and according to sex; and the eighteen resulting group summaries bear almost no resemblance to each other, each color combination having been given each of the six ranks at least once. That there is practically no agreement among young children as to their reactions to combinations of colors would seem, then, to be indicated. Now, again, whether this lack of a common basis is due merely to undeveloped, unorganized character, or is due to the limited amount and varying nature of the social control by suggestion, would make an interesting problem. That in *some way* it is due to immaturity, is, however, established by a comparison with the results from the college students.

TABLE VII

Sophomore Girls and Boys (125)								
	1	2	3	4	5	6	T. V.	R.
Red-green	10	17	27	25	21	25	480	5
Red-blue	30	26	22	20	15	12	375	1
Orange-yellow	18	16	27	28	27	9	442	4
Orange-green	8	15	16	16	31	39	539	6
Yellow-violet	41	19	15	14	18	18	378	2
Blue-violet	18	32	18	22	13	22	421	3

That the preferences are again more uniform and more

definite than with children is exhibited, *e. g.*, by the difference between T. V. 375 for red-blue (ranked first) and 539 for orange-green (ranked sixth).

In the matter of similarities between actual choices made by children and college students, perhaps the only correlations of importance are in ranking the orange-green lowest and the red-blue at or near the top.

Sex differences among the sophomores are interesting.

Table VIII

Sophomore Girls (91)			Sophomore Boys (34)		
	T. V.	R.		T. V.	R.
Red-green	365	5	Red-green	115	3
Red-blue	295	3	Red-blue	80	1
Orange-yellow	328	4	Orange-yellow	114	2
Orange-green	421	6	Orange-green	118	4
Yellow-violet	227	1	Yellow-violet	151	6
Blue-violet	285	2	Blue-violet	136	5[1]

The girls show a preference for the yellow-violet combination with blue-violet and red-blue rated good; while the boys show a strikingly consistent preference for the red-blue, rating the other two just mentioned low. And this difference between the sexes is clear in the summaries for each of the three sections into which the students were divided.

That the sex differences here are not similar to those at the kindergarten age (Table VI.) would logically follow from the general dissimilarity, noted above, in ranking by boys and girls together at the two ages (Tables V. and VII.).[2]

III

In the experiment on tone-intervals the favorable or unfavorable reaction as expressed by the child for each

[1] Jastrow got a somewhat different result with his adults. Using a greater range of color combinations he found that women place light red-light green first, red-green second, red-light green third; men place red-blue first, blue-violet second.

[2] The studies of Dobbie and of Wundt on children's preferences in color combinations were concerned with a somewhat different problem. Dobbie shows that "the combinations with minimum difference (spectral order) [between successive colors] has a decided preference." This confirms Wundt's finding that his two children found most pleasing those combinations which showed spectral proximity and order.

interval sounded was recorded by a positive (+) or negative (−) sign respectively.

TABLE IX

	Kindergarten Girls (99)		Kindergarten Boys (99)		Both (198)	
	+	−	+	−	+	−
c—e........	97	2	96	3	193	5
c—b........	69	30	63	36	132	66
c—g........	91	8	93	6	184	14
c—c........	89	10	87	12	176	22
c—d^b.......	53	46	58	41	111	87

Compare this with the table for the sophomores.

TABLE X

	Sophomore Girls (63)		Sophomore Boys (26)		Both (89)	
	+	−	+	−	+	−
c—e........	59	4	25	1	84	5
c—b........	—	63	—	26	—	89
c—g........	60	3	21	5	81	8
c—c........	60	3	25	1	85	4
c—d^b.......	1	62	—	26	1	88

Again, the striking difference between the two ages of subjects is one of definiteness and reliability of results. As was to be expected from the simplicity of the intervals the college students reacted almost unanimously. The intervals are well recognized as pleasant or unpleasant by adults in general. But the children's reactions, with such great variability in their likes and dislikes for such intervals, may be a bit unexpected.

What is to be made of this? (*A*) Shall we infer that the comparatively unorganized character of child-nature, the absence as yet of æsthetic sense, is the complete explanation: that all that is necessary to get results as in Table X. from subjects of Table IX. is growth and natural mental development? Such an interpretation would seem to imply that by their very physical nature and the relationship of that to the physiological nature of the adult human subjects, intervals c–b and c–d^b are and must be such as to arouse unpleasant

reactions, c–e, c–g, and c–c, pleasant reactions. This more objective and absolute interpretation must take into consideration the character of the component physical vibrations, and whatever is to be known about the physiological effects of differing combinations of vibration frequencies. (*B*) Or, shall we insist that something of an æsthetic sense is already present, but that it is modifiable, plastic; and that between the kindergarten and the college ages, come the multiplicity of social suggestions that tend to make all individuals of a society more or less alike in their standards of what is and what is not likable? To put this point more concretely: it is conceivable, at least, that if put in a social community in which the intervals c–b and c–d^b were handled as pleasant intervals and the musical forms of such a community based upon the pleasantness of these and the unpleasantness of the c–e, c–g, and c–c intervals, the child would develop a set of preferences totally different from those shown by the college students above.

And, of course, the (*A*) and (*B*) interpretations are quite relevant to the matter of differences in color preferences as set forth in the earlier paragraphs of this paper.

Summary

In the experiments on single colors and on tone intervals but not on color combinations there is a rough qualitative agreement between the order of judgments made by the two ages.

With single colors the two sexes at both ages show a similarity in preferring certain colors but a real difference in the order in which they are preferred. With the exception of red—which seems to involve an unknown disturbing factor—age makes little difference. With color combinations sex comparisons are valueless with the children, but with the sophomores they indicate preference for red-blue by boys and yellow-violet by girls. With tone intervals there are no sex differences brought out.

The outstanding results of all three sets of experiments is the clear demonstration of a decidedly greater variability

and unreliability (for group purposes, at least) of æsthetic judgments by children of kindergarten age as compared with those by later adolescents and adults. What this may be due to has been suggested as a fundamental problem.

Pedagogical applications would seem to follow. In general, it would appear that kindergarten methods which are now necessarily based more or less on *a priori* foundations and on general observations might benefit from critical studies of how the child does actually react to æsthetic (and other) stimuli.

THE RELATION BETWEEN LEARNING AND RETENTION AND AMOUNT TO BE LEARNED[1]

BY V. A. C. HENMON

University of Wisconsin

In the experimental study of memory the first problem taken up was the relation between the length of series of nonsense-syllables and the number of repetitions required for learning them. Does the difficulty of learning increase proportionally with an increase in the number of syllables, more slowly or more rapidly? Ebbinghaus[2] in the widely quoted results found that while 7 syllables could be recalled with one presentation, 12 syllables required 16.6 repetitions, 16 syllables 30 repetitions, 24 syllables 44 repetitions, and 36 syllables 55 repetitions. In other words, the number of repetitions increases at first with very great rapidity and then less rapidly, the increase in repetitions being relatively much greater than the increase in the number of syllables. Binet and Henri[3] found a similar tendency in the learning of num-

[1] Paper read before the American Psychological Association at the Washington meeting, 1911. Publication has been delayed in the expectation that further work could be done on the problem. In view of the full report of the paper by Lyon, D. O., *J. of Educ. Psychol.*, 1914, 5, pp. 7–9, and the review by Pieron, H., *Année Psychol.*, Vol. XIX., 1913, pp. 119–120, it seems best to publish the paper as read.

[2] Ebbinghaus, H., 'Über das Gedächtnis,' p. 64. 'Grundzüge der Psychologie, pp. 650–651. The figures for these experiments have been often and variously reported:

Ebbinghaus (Grundzüge)	Syllables	6	12	16	36
	Repetitions	1	14–16	30	55
James	Syllables	7	12	24	26
	Repetitions	1	16	44	55
Meumann	Syllables	7	12	16	24–36
	Repetitions	1	16.6	30	44–65
Claparède	Syllables		12	24	36
	Repetitions		16.5	44	55
Myers	Syllables	7	12–16	24	36
	Repetitions	1	16.6–30	44	55

[3] Binet, A., & Henri, V., 'La Memoire des Mots,' *Année Psychol.*, 1895, 1, p. 12.

bers. The increase in time for learning with the increase in length of series was as follows:

10 numbers	17 secs.
15 "	75 "
20 "	135 "
25 "	180 "
30 "	260 "
50 "	420 "
100 "	1,500 "
200 "	4,520 "

The explanation of this result is sought by Ebbinghaus in the narrowness of the span of consciousness and in retroactive inhibition. The narrowness of the span of consciousness accounts for the increase in repetitions with twelve syllables over seven. Retroactive inhibition accounts for the increase with the longer series. This retroactive inhibition is exerted by each individual association not only upon those immediately preceding it, but to some extent on all those more remote as well. The effect, therefore, is greater the greater the number of inhibiting and inhibited members. Offner[1] suggests as a partial explanation that the greater the number of members in a series the less attention given to each member, apparently assuming that the amount of attention to any series is a constant quantity and that in a long series it is merely more widely distributed; hence, the greater number of repetitions for learning. He finds a second factor in the influence of forgetting which begins immediately and has progressed so much further in a long series than in a short one that it requires numerous repetitions to bring it back to the level of a short series. Myers[2] attributes the result to fatigue and the lesser concentration of attention with the longer series.

No systematic inquiry has been made to test the validity of Ebbinghaus's law and his data are quoted in all the books. This seems curious in view of the fact that Ebbinghaus expressly states that his results are tentative and that he

[1] Offner, M., 'Das Gedächtnis,' Berlin, 1909, p. 54–55.

[2] Myers, C. S., 'Text-Book of Experimental Psychology,' London, 1909, pp. 157–158.

does not regard them as very reliable. More recent investigations, however, so far as they have gone, have failed to confirm them.

Radossawljewitch[1] noted in his experiments on the rate of forgetting that the change from 8 to 12 or to 16 syllable series did not demand a great increase in the number of repetitions, in fact, frequently 16 syllables were memorized with fewer repetitions than 8 or 12 syllables. To test the matter more fully Meumann learned 5 series each of 8 and 12 syllables and 2 series each of 16, 18, 24 and 36 syllables with the following results:

No. of Syllables	Repetitions	Differences
8	5.2	
12	10.4	5.2
16	17.0	6.6
18	21.5	4.5
24	30.0	8.5
36	32.5	2.5

These results are the reverse of those obtained by Ebbinghaus. There is a relative decrease in repetition with an increase in the number of syllables. Meumann[2] holds that this is what might be expected. An increase in amount of work to be done, if it is not too great, makes little difference when once the initial disinclination or inertia is overcome, when adaptation of attention is secured, when the associative processes have been aroused and a general adjustment to the work is once attained. All of these formal conditions of learning should be effective for a series no matter what their length within the limits of fatigue. Hence it is reasonable to expect a relative decrease in energy required for learning with an increase in amount to be learned.

The relation between amount to be learned and retention as distinguished from learning is not clearly made out. Ebbinghaus found that the longer the series the more firmly it is impressed; hence after 24 hours both the absolute and

[1] Radossawljewitch, P., 'Das Behalten und Vergessen bei Kindern und Erwachsenen,' Leipzig, 1906.

[2] Meumann, E., 'Über Ökonomie und Technik des Gedächtnisses,' Leipzig, 1908.

relative economy is greatest with the longest series. The economy after twenty-four hours was as follows:

No. of Syllables	Per Cent. Saved.
12	33
24	49
36	58

Binet and Henri[1] and Pentchew[2] confirmed this result. Radossawljewitch[3] also notes that the shorter series were more quickly forgotten than the longer series except for the intervals of 5 minutes and 20 minutes.

Müller and Pilzecker and Reuther, however, claim that while the absolute amount retained increases with the increase in the length of the series, the relative amount decreases.

Ebbinghaus' results suggest that economy in mental work would result from a subdivision of it into small amounts. This flies in the face of common experience and the results of the experiments on learning by wholes and by parts. From Meumann and Radossawljewitch's results one would conclude that it is more economical to take large amounts at a time, the larger the amounts, the better within the limits of fatigue and available psychophysical energy.

Ebbinghaus' experiments are on one subject only. Meumann's results are also on one subject and the number of cases small, the figures for the longer series being averages from two series. Hence, it seemed worth while to continue the investigation on a number of individuals with nonsense-syllables and also with sense material, poetry and prose.

Experiments and Results

1. *Nonsense Syllables.*—The first experiment was designed to test Ebbinghaus and Meumann's results in the learning of nonsense-syllables. Eight series each of 10, 12, 14, 16, 18, 20, 24 and 30 syllables were prepared for exposure with the Ranschburg apparatus. These were presented to the subjects

[1] *Op. cit.*

[2] Pentchew, C., 'Untersuchungen zur Ökonomie und Technik des Lernens,' *Arch. für d. ges. Psychol.*, I, 1903.

[3] *Op. cit.*

at a rate which permitted the exposure of each syllable for 1¼ seconds. Two subjects took part in the experiments. Ordinarily, the subject worked for one hour at the same time a day and relearned after twenty-four hours.

The results of the experiment, for Subject *H*, appear in Table I. It gives the average number of repetitions required for learning, based on 8 experiments and for relearning based on 4–6 cases, the mean variations, and the per cent. of repetitions saved after twenty-four hours.

TABLE I.

	No. of Syllables							
	10	12	14	16	18	20	24	30
Repetitions for learning	6.9	8.1	8.3	8.6	10.6	13.6	12.9	20.0
M. V.	.9	1.5	1.1	1.6	1.6	2.4	1.5	3.3
Repetitions for relearning	3.8	4.5	4.8	4.4	5.5	6.5	5.7	9.5
M. V.	.4	.5	.4	.6	.5	.8	.7	.5
Per cent. saved after twenty-four hours	45	44	42	49	48	52	56	53

Table II. gives the results for Subject *C* for learning, the averages being from 3 experiments.

TABLE II.

	No. of Syllables							
	10	12	14	16	18	20	24	30
Repetitions for learning	12.7	14.0	15.3	15.0	16.3	19.0	16.0	25.7
M. V.	1.1	1.3	1.6	0.0	1.8	1.3	2.7	3.6

Fig. 1 indicates graphically the curve of increase in repetitions and for purposes of comparison the curves obtained by Ebbinghaus (E) and Meumann (M) are included.

The results of the experiments diverge widely from those of Ebbinghaus and Binet and are in fairly close accord with those of Meumann and Radossawljewitch. There is a relative decrease in the number of repetitions as the number of syllables increases. Particularly noteworthy is the fact that the number of repetitions for the series from 10 to 18 is practically constant. The results are even more striking than those of Meumann in showing the relative economy with

the longer series. It seems clear that Ebbinghaus' results do not represent a general tendency.

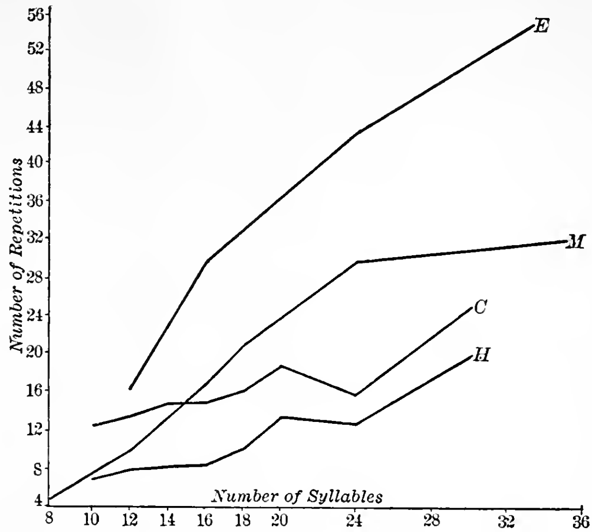


Fig. 1.

As is to be expected the saving in repetitions after twenty-four hours will not show the marked differences which Ebbinghaus obtained. Ebbinghaus found a saving of 33 per cent. in repetitions with 12 syllables after twenty-four hours, of 49 per cent. with 24 syllables, of 58 per cent. with 36 syllables. In other words, the force of association increases progressively with the increase in the number of syllables. In the results with Subject *H*, the saving after twenty-four hours is also slightly greater with the longer series. Both from the point of view of repetitions required for learning and for relearning after twenty-four hours the economy is greatest with the longer series.

Weber working in Meumann's laboratory reports that the law holds for subjects untrained in learning syllables. A repetition of the experiments on six untrained subjects with 10, 12, 14, 16, 18, 20 syllables fails to confirm this result. The data appear in Table III. (The figures in parentheses indicate the number of experiments.)

TABLE III.

	10	12	14	16	18	20
A (3)	17.7	28.7	22.7	22.7	22.3	38.3
B (2)	14.0	20.0	19.0	23.5		
C (2)	11.0	18.0	23.5	25.0		
D (1)	20.0	25.0	27.0	31.0		
E (2)	9.0	7.5	9.5	16.0		
F (1)	9.0	10.0	11.0	26.0		
Av.	13.4	18.2	18.8	24.0	22.3	38.3

None of these subjects had ever learned nonsense-syllables. To eliminate the influence of practice the order of the series varied with each subject, which accounts for the irregularities in the results, *e. g.*, *A* learned 12 syllables first, *F*, 16 syllables, etc.

2. *Poetry.*—The second set of experiments consisted in memorizing by the 'whole' method 1, 2, 3, 4 and 5 stanzas of 'In Memoriam,' a record being kept of the number of repetitions required for learning and for relearning after twenty-four hours. The results on three subjects, giving the repetitions for learning (L.), for relearning after twenty-four hours (R.), the per cent. saved and the mean variations appear in Table IV. The figures in parentheses give the number of experiments with each group of stanzas.

TABLE IV.

Stanzas	H (10)					D (10)					P. (5)					Average		
	L.	M. V.	R.	M. V.	Economy, Per Cent.	L.	M. V.	R.	M. V.	Economy, Per Cent.	L.	M. V.	R.	M. V.	Economy, Per Cent.	L.	R.	Economy, Per Cent.
1	3.3	.6	1.3	.4	61	3.5	.7	1.1	.4	69	3.8	.6	2.0	.4	47	3.5	1.5	57
2	5.3	.7	2.2	.3	59	6.1	.9	1.8	.5	70	7.4	1.5	3.0	.4	59	6.3	2.3	63
3	6.7	.9	2.3	.6	66	8.9	1.3	1.4	.9	84	10.2	1.8	3.8	.8	63	8.6	2.5	71
4	7.5	1.1	2.5	.5	67	10.7	1.9	1.9	.5	82	11.8	1.8	3.8	.6	68	10.0	2.7	73
5	8.6	1.2	2.6	.6	69	14.0	1.8	2.3	1.6	80	14.0	2.8	4.6	.7	67	12.2	3.2	74

The increase in the number of repetitions with the increase in amount is relatively less than the increase in the number of lines or stanzas. If the increase was proportional to the amount the number of repetitions would be 3.5, 7.0, 10.5,

14.0 and 17.5 instead of which the series is 3.5, 6.3, 8.6, 10.0, and 12.2. There is thus a relative economy with the larger amounts. The economy in relearning after twenty-four hours is greater with the larger amounts and is relatively greater with poetry than with nonsense syllables.

'In Memoriam' is rather difficult material for memorizing and the rate of increase is relatively rapid. One practiced subject with 1, 2, 3, 4, stanzas of 'Childe Harold' gave a less rapid increase.

Stanzas	L.	R.	Per Cent. Saved
1	8.3 Rep.	4.8 Rep.	40
2	9.3 "	5.5 "	40
3	10.0 "	5.4 "	46
4	9.1 "	4.8 "	47

A test of nine unpracticed subjects gave the following results:

No. of stanzas......	1	2	3	4	5
Repetitions........	4.6	9.7	11.8	15.9	15.3

3. *Prose.*—For the experiments on learning prose, 100-word, 200-word, and 300-word passages were selected from the essays of Huxley and Matthew Arnold. 60 selections from Huxley were made, and 60 from Arnold, 20 passages of each length. One practiced subject learned 54 of these selections, 18 from each group, and recorded the number of repetitions required for learning. Eight of each of the selections were relearned after twenty-four hours and 10 of each group were relearned after seven months. The rate of reading in learning was made as uniform as possible. The results are indicated in Table V.

TABLE V.

	100 Words	200 Words	300 Words
Learning Huxley passages..............	7.4 Rep.	7.9 Rep.	7.4 Rep.
Relearning after twenty-four hours.......	4.0	3.0	2.8
Per Cent. saved.......................	46%	62%	62%
Learning Arnold passages...............	5.5 Rep.	6.7 Rep.	6.7 Rep.
Relearning after seven months...........	4.3	5.0	4.3
Per Cent. saved.......................	21%	25%	35%

There is an approximate constancy in the number of repetitions for the passages irrespective of their length. The averages for the 18 passages of each length are:

	100 words	200 words	300 words
Repetitions......	6.1	7.3	7.0

It may be remarked that a double fatigue order in learning was observed so as to eliminate inequalities in practice effect.

Experiments which have been made on the distribution of repetitions and on learning by wholes and by parts would lead one to infer that Ebbinghaus's law could not well represent a general tendency. My results show the inference to be correct. Retroactive inhibition, forgetting, the distribution of attention and fatigue which have been invoked to explain Ebbinghaus' results are more than compensated for by perseveration, by the time permitted for the setting of associations, and by the greater effort of attention which the longer series calls forth.